

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581
Sequence: 1 MKVLMALVTFELAGCQAKV.....VEKQAVNGTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n.model -DEV=xip
-O=/cgn2.1/USPFO/spool/US09827854/runat_11032003.101610.27486/app-query.fasta.1.3576
-DB=genEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 -EGN1.1_13008 -etrunat_11032003.101610.27486 -NCPU=6 -ICPU=3
-NO_XLUPX -NO_MMAP -LARGEQUERY -NEG1SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hc:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	6	BD004279 Apo E hum
2	1573	99.5	1110	6	E00359 cDNA coding
3	1573	99.5	1110	6	E00823 DNA sequence
4	1573	99.5	1147	6	AX302545 Sequence
5	1573	99.5	1156	6	BD004278 Apo E hum
6	1573	99.5	1156	9	H00396 Homo sapien
7	1573	99.5	1186	9	BC003557 Homo sapi
8	1566	99.1	1110	6	E08423 DNA coding
9	1565	99.0	1156	6	BD004277 Apo E hum
10	1564	98.9	1157	6	AX333278 Sequence
11	1564	98.9	1157	6	AX409597 Sequence
12	1564	98.9	1157	6	AX409597 Sequence
13	1564	98.9	1157	6	AX409597 Sequence
14	1459	98.9	1157	9	H00406E HMAPOE
15	1387	87.7	5515	9	H00406E HMAPOE
16	1384.5	87.6	5491	9	AF261279 Homo sapi
17	1384.5	87.6	41907	6	AX358722 Sequence
18	1384.5	87.6	41907	9	AF050154 Homo sapi
19	1384.5	87.6	107567	9	AC011481 Homo sapi
20	1358	85.9	5413	9	AF261280 Homo sapi
21	1275	80.6	4762	9	BABAPOE
22	1274	80.6	208239	2	AC021988
23	1266	80.1	1138	4	AF303830
24	1159	73.3	718	9	AF20049753
25	1158	73.2	1060	4	RABAPOLP
26	1157	73.2	1045	10	MUSAPOE
27	1157	73.2	1104	10	BC028816
28	1155	73.1	718	9	AF20050053
29	1155	73.1	718	9	AF20050053
30	1153.5	73.0	1154	4	BTAPOEHR
31	1151.5	72.8	965	6	AX384545
32	1151.5	72.8	1108	4	BTAPOLPE
33	1151.5	72.8	5617	6	AX384541
34	1151.5	72.8	6026	6	AX384539
35	1151	72.8	1122	4	AF20050653
36	1151	72.8	1122	4	SSAPOE
37	1146	72.5	1126	6	ARI64342
38	1146	72.5	1126	6	AR205885
39	1102	69.7	951	10	S76779
40	1080	68.3	959	10	MUSAPOE
41	1064.5	67.3	1069	10	RAIPAOE
42	1037	65.6	228698	2	AC127479
43	1037	65.6	237653	2	AC073760
44	1011	63.9	4856	10	MUSAPOE
45	1010	63.9	4267	6	ARI64387

RESULT 1

ALIGNMENTS

[illegible][illegible]

source 1..1110
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN

Alignment Scores:
Pred. No.: 1.95e-77 Length: 1110
Score: 1573.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.49% Indels: 0
DB: Gaps: 0

US-09-827-854-14 (1-317) x E00359 (1-1110)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGGTTCTGTGGCTCTGCTGTGCATTCCTGGCAGATGCCAGGCCAAGGTG 74

QY 21 GtGlnAlaValGtThcGluProGluLeuArgGlnGlnThcGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGGCCAGCAGACGAGTGGCAGAGC 134

QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCCAGCGCTGGGAACCTGGCATTGGCTTTGGATTACTGGCTGGGGTGCAGACA 194

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTGTGACAGGTGACAGAGAGGTGTCTCAGCTCCAGAGTCACCCAGGAACCTGAGGGCG 254

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 255 CTGATGAGCAGACCATGATGAGAGCTTGAGGCTTACAAATCGGAAGTGGAGAACACTG 314

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 315 ACCCGGTGGCGGAGAGACGGCGGCGCTGCCAAGAGACTGCAGGGCGGCGAGCGC 374

QY 121 ArgLeuGlyAlaAspMetClnAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 375 CGCGTGGGGCGGACATGAGAGACGTGTGCGCGCCCTGTGTACAGTACCGCGCGAGGTG 434

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGCCAGAGACACCGAGAGCTGGGGGTGGCTGCCCTCCACCTGGCC 494

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 495 AAGCTGCGTAAGCGGCTCTCCGCAATGCCATGACCTGCACAAACGCTGGCAGTATAC 554

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
DB 555 CAGGCGGGGGCGCGGAGGGCGCCGAGCGGCTCAGGGCCATCCGAGAGCGCTGGGG 614

QY 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGTGTGAGACAGAGCGCGCTGGCGGCTGTGGCTCCCTGGCGGCGAGCGG 674

QY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
DB 675 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGCGATGGAGAGATGGGGC 734

QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGluValArgAlaLys 260
DB 735 AACCCGAGCCCGGACCGCTGGAGCAGGTGAAGGAGAGGTGGCGAGTGGCGGCCCAAG 794

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaLysLeuLys 280
DB 795 CTGGAGAGAGAGCCCGACAGATACGCTGCAGGCGCGAGGCTTCCAGGCGCGCTCAAG 854

QY 281 SerTrpPheGluProLeuValGlnLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300

DB 855 ACCTGGTTGACCCCTGGTGGAGACAGCAGCGGCATGGGGCGGCTGGTGGAGAG 914

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTGCAGGCTGCCGTGGGACACAGCGCCCGCTGTGTCCAGCAGCATATCAC 965

RESUT 3
E00823
LOCUS E00823 1110 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.
ACCESSION E00823
VERSION E00823.1 GI:2169084
KEYWORDS JP 1986096997-A/1.
SOURCE unidentifed.
ORGANISM unclassified.

REFERENCE
AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL Patent: JP 1986096997-A I 15-MAY-1986;
MITSUBISHI CHEM IND LTD
OS Human {Homo sapiens}
PN JP 1986096997-A/1
PD 15-MAY-1986
PF 16-OCT-1984 JP 1984216987
PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO
PC C12P21/00,A61K35/74,A61K37/04,C12N15/00,C12P21/00,C12R1:19),
PC (C12N15/00,
PC C12R1:19);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
*source: tissue type=liver;
key location/Qualifiers

FT 3'UTR 1..14
FT sig-peptide 15..68
FT /product='human apolipoprotein E signal FT
FT CDS 69..968
FT peptide'
FT /product='human apolipoprotein E' FT 3'UTR

FEATURES
source 969..1110. /product='human apolipoprotein E' FT
location/Qualifiers
1..1110
/organism='unidentifed'
/db_xref="taxon:32644"

BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN

Alignment Scores:
Pred. No.: 1.95e-77 Length: 1110
Score: 1573.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.49% Indels: 0
DB: Gaps: 0

US-09-827-854-14 (1-317) x E00823 (1-1110)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGGTTCTGTGGCTCTGCTGTGCATTCCTGGCAGATGCCAGGCCAAGGTG 74

QY 21 GtGlnAlaValGtThcGluProGluLeuArgGlnGlnThcGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGGCCAGCAGACGAGTGGCAGAGC 134

QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCCAGCGCTGGGAACCTGGCATTGGCTTTGGATTACTGGCTGGGGTGCAGACA 194

QY	61	LeuSerGluInuValGlnInuGluInuLeuSerSerGlnValThrGlnInuLeuArgAla	80
Db	195	CTGTCTGACAGAGTGGAGAGAGAGCTGTCAAGTCCACAGATCACCCAGGAACGTGAGAGGG	254
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnInuGlnLeu	100
Db	255	CTGATGACGAGAACCATGAAAGAGTTGAAGGCCCTACAAATCGGAATCGAGGAACAATG	314
QY	101	ThrProValAlaGluInuThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	315	ACCCCGGTGGCGAGAGACGCGGGACGGCTGTCCAAAGAGCTGCAGCGGGCGAGGCC	374
QY	121	ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	375	CGGCTGGGGCGGACATGTGAGAGACGtGTGGGGCGGCTGTGTGACATGACGGGGCGAGGTG	434
QY	141	GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCCATCTCTGGCCAGAGACACCGAGGAGCTGGGGGTGGCTCCCTCCACCTGGCC	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspLysGlnLysArgLeuAlaValTyr	180
Db	495	AAGCTCCGTACAGGGCTCTCTCCGCCATGCCGATGACCTCAGAAAGCGCTGGCAGGTGTC	554
QY	181	GlnAlaGlyAlaArgGluInuGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	555	CAGGCGGGGGCCCGCGAGGGCGCCGAGCGCGCTCAGCGCCATCCGAGCGCCCTGGGG	614
QY	201	ProLeuValGluGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	615	CCCCTGTGTGAACAAGGCCGCGTGGGGGGCGCCACTGTGGCTCCCTGGCGGGCAGCGG	674
QY	221	LeuGlnGluArgAlaGlnAlaThrTyrGlyGluArgLeuArgAlaArgMetGluLysMetGly	240
Db	675	CTACAGAGAGGGGGCCCAAGGCTGTGGGGCGAGCGGCTGTGGCGCGGATGGAGGAGATGGGG	734
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	735	AACCGGACCCCGAGCGCCCTGTGACGAGGTGAAGAGAGACAGTGGCGAGGTGGCGGCAAG	794
QY	261	IeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
Db	795	CTGGAGAGACAGAGCCAGACAGATACGCCCTGCAGGCCGAGGCCCTTCAGAGCCCGCTCAAG	854
QY	281	SerTyrPheGluProLeuValGluLysMetGlnArgGlnTyrPheAlaGlyLeuValGluLys	300
Db	855	AGCTGGTTCGACCCCTGTGTGAACAATGACAGCGCAATGTGGCGCGGCTGTGGAGAG	914
QY	301	ValGlnAlaAlaValGlyLysSerHisAlaAlaProValProSerAspAsnHis	317
Db	915	GTCAGAGCTGCCGTGGGCACACGCGCGCTGTGCCACGAGACAAATCAC	965
RESULT 4			
LOCUS	AX302545	1147 bp	DNA linear PAT 30-NOV-2001
DEFINITION	Sequence 63 from Patent W00175177.		
ACCESSION	AX302545		
VERSION	AX302545.1		
KEYWORDS	GI:17383082		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Molin, P.J., Sherman-Baust, C.A., Fizer, E.S. and Hough, C.D.		
TITLE	Tumor markers in ovarian cancer		
JOURNAL	Patent: WO 0175177-A 63 11-OCT-2001.		
FEATURES	THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
Source	Location/Qualifiers		
	1..1147		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

BASE COUNT	210 a	365 c	425 g	147 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	2.02e-77			Length:	1147
Score:	1573.00			Matches:	316
Percent Similarity:	99.68%			Conservative:	0
Best Local Similarity:	99.68%			Mismatches:	1
Query Match:	99.49%			Indels:	0
DB:	6			Gaps:	0
US-09-827-854-14 (1-317) x AX302545 (1-1147)					
QY	1	MctLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlySGlnAlaIleVal	20		
DB	46	ATGAAGGTTCTGTGGGCTGCGTTGCTGGTCAATTCTCGCAGAGATGCCAGGCCAAGTG	105		
QY	21	GIUGlnAlaValaGIuThrGIuProGIuProGIuLeuAArgGIuInrGIuTrpGIuSer	40		
DB	106	GAGCAAGCGGTGAGACAGACAGCCGGAGCCGAGCTGCCAGCAGCAGCAAGTGGCAGAGC	165		
QY	41	GIyGlnAArgTrpGIuLeuAlaIleuGIyAArgPheTrpAspTrpLeuAArgTrpValGIuThr	60		
DB	166	GGCCAGCCCTGGGAAGTGGCACTGGCTGGCTTTGGATTACCTGGCCTGGGTGCACACA	225		
QY	61	LeuSerGIuInrValGIuGIuLeuLeuSerSerGIuValThrGIuInrGIuLeuAArgAla	80		
DB	226	CTGTCTGAGCAGATGTCAGGAGAGAGCTGCTCAGCTCCACAGTCCACCCAGAACTGAGGGCG	285		
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTrpLysSerGIuLeuGIuInrGIuLeu	100		
DB	286	CTGATGGACGAGACCATGAAGAGATTGAAGSCTTACAATAATGGAACTGAGAAACAAGTGG	345		
QY	101	ThrProValAlaGIuGIuThrThrAlaAArgLeuSerLysGIuLeuGIuAlaIleGIuAla	120		
DB	346	ACCCCGGTGGGGAGAGAACCCGGGCAAGCGCTGTCCAAAGAGCTGCAGCGCGCGCAGGCC	405		
QY	121	AArgLeuGIyAlaAspMetGIuAspValAArgGIyAArgLeuValGIuTrpAArgGIyGIuVal	140		
DB	406	CGCGTGGGCGCGGAGATGAGAGAGCTGTGGCGCCGCTGGTGCAGTACCCGGCGAGGTG	465		
QY	141	GlnAlaMetLeuGIyGIuSerThrGIuGIuLeuAArgValAArgLeuAlaSerHisLeuAArg	160		
DB	466	CAGGCATGCTGTGGCCAAAGACACCGAGAGAGCTGGCGGTGGCCCTCCCTCCCACTGGCCG	525		
QY	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspIleuGIuLysAArgLeuAlaValTyr	180		
DB	526	AAGCTGCGCTAAGCGGCTCTCTCGGATGTCGATACCTGCAGGAAGCCCTGGCAGTGTAC	585		
QY	181	GlnAlaGIyAlaAArgGIuGIyAlaGIuAArgGIyLeuSerAlaIleAArgGIuAArgLeuGIy	200		
DB	586	CAGCGCGGGGCGCGGAGAGCGCGCGAGCGCGCTTCAGCCGCACTCCGCGAGCCGCTGGGG	645		
QY	201	ProLeuValGIuGIuGIyAArgValAArgAlaAlaThrValGIySerLeuAlaGIyGIuPro	220		
DB	646	CCCTGCTGTGAACAGGGCGCGGTGGGGCGCCCACTGTGGGCTCTCCGTGGCGGCCACGCCG	705		
QY	221	LeuGIuGIuAArgAlaGIuAlaIleTrpGIyGIuAArgLeuAArgAlaAArgMetGIuGIuMetGIy	240		
DB	706	CTACAGGAGCGGGCCAGAGCCTGTGGGGCGAGGCGCTGGCGCCGCGCGAGAGGAGATGGCG	765		
QY	241	SerAArgTrpAArgAspAArgLeuAspGIuValLysGIuInrValAlaGIuValAArgAlaLys	260		
DB	826	CTGGAGGAGCAGGCCCAAGACATACGCTGCAGGCGCAGAGGCTTTCACAGCCCTTCACAG	885		
QY	261	LeuGIuGIuGIuAlaGIuGIuGIuLysAArgLeuGIuAlaGIuAlaPheGIuAlaAArgLeuLys	280		
DB	886	AGCTGGTTCGAGCCCTGTGTGAAGACATGACGCGCAAGTGGCGGGGCTGTGTGAAG	945		

Qy 301 ValGlnAlaIaValaGlyThrSerAlaIaProValProSerAspAsnHis 317
 |||
 Db 946 GTGCAGGCTGGCGTGGACACGAGCGCCCTGTGTGCCAGGACAAATCAC 996

RESULT 5
 BD004278 1156 bp DNA linear PAT 31-JAN-2002
 LOCUS Apo E humanized mammal.
 DEFINITION BD004278
 ACCESSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1156)
 Fujita,S., Hamanaka,H.,
 Apo E humanized mammal
 Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PE 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
 A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
 C12N15/09//C07K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

FEATURES
 source 1..1156 location/Qualifiers
 FT CDS (61)..(1011).
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.03e-77 Length: 1156
 Score: 1573.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.49% Indels: 0
 DB: Gaps: 0

US-09-827-854-14 (1-317) x BD004278 (1-1156)

Qy 1 MetLysValLeuTRPAlaIaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 Db 61 ATGAAGTTCTCTGTGGCTGGCTGCTGGTCACATTCTCTGGCGAGATGCCAGGCGAGGTG 120

Qy 21 GlnGlnAlaValaGlyThrGluProGluProGluLeuArgGlnGlnThrGluTrpLysIns 40
 |||
 Db 121 GAGCAAGGCGGTGAGACAGAGCGGAGCGCGGAGCTGCCGACACACGAGAGTGGAGAGC 180

Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||
 Db 181 GGCCAGCGCTGGAACTGGCACTGGCTGCTTTGGATTACCTCGCTGGGTGCAGACA 240

Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgLys 80
 |||
 Db 241 CTGTCTGAGCGGTGAGAGAGGCTGCTGCTGCTCCAGGTCACCCAGGAATGAGGCGC 300

Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
 |||
 Db 301 CTGATGAGACAGACCATGAGAGGTGAAGGCTTCAAAATGCGACTGGAGAACAACTG 360

Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaLysVal 120
 |||
 Db 361 ACCCGGTGGCGGAGAGACGCGGCGCTGTCTCAAGAGAGCTGACAGCGCGCGCAGCC 420

Qy 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140
 |||
 Db 421 CGGCTGGGCGCGGATAGAGAGCTGTGGCGCCCTGTGTGCAATACCGCGGAGAGGTG 480

Qy 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||
 Db 481 CAGGCGATGCTGCGCCAGAGACACCGAGAGCGCGGCTGCTGCTCCACCTGGCGC 540

Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 |||
 Db 541 AAGCTGCTTAAGCGGCTCTCCGCGATCCGATGACCTGCGAGAACCGCTGGCACTGTAC 600

Qy 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||
 Db 601 CAGCGCGGCGCCCGGAGAGCGCCGAGCGGCTTACAGCGCATTCGGAGAGCGCTGGGG 660

Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyPro 220
 |||
 Db 661 CCCCTGTGTGAACAGAGCGCGGTGCGGCGCGCCACTGTGTGGCTCTGCGCGCCAGCCG 720

Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
 |||
 Db 721 CTACAGAGAGCGGCGCCAGGCTGTGGGCGAGCGGCTGCGCGGATGAGAGATGGGC 780

Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValaGluValaArgAlaLys 260
 |||
 Db 781 AGCCGAGACCCCGGAGACCGCTGTGACAGGTGAAGAGAGAGTGGCGGCGCCAAAG 840

Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||
 Db 841 CTGCAGAGAGAGCGCCACAGATACGCTGCAGCGCGGAGGCTTCCAGCGCGGCTCAAG 900

Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||
 Db 901 AGCTGTTCGAGCGCCCTGTGTGAAGACATGACGCCCTGAGCGGCTGTGTGAGAG 960

Qy 301 ValGlnAlaIaValaGlyThrSerAlaIaProValProSerAspAsnHis 317
 |||
 Db 961 GTGCAGGCTGGCGTGGACACGAGCGCCCTGTGTGCCAGGACAAATCAC 1011

RESULT 6
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS Homo sapiens preapolioprotein E (APOE) mRNA, complete cds.
 DEFINITION K00396
 ACCESSION K00396.1 GI:178850
 KEYWORDS apolioprotein; apolioprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 355 to 1156)
 Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,
 Iofquist-Kahl,F., Karathanasis,S.K. and zannis,V.I.
 Identification and DNA sequence of a human apolioprotein E cDNA
 clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 2 (bases 250 to 777)
 WALLIS,S.C., ROGNE,S., GILL,L., MARKHAM,A., EDGE,M., WOODS,D.,
 WILLIAMSON,K. and HUMPHRIES,S.
 The isolation of cDNA clones for human apolioprotein E and the
 detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 3 (bases 1 to 1156)
 ZANNIS,V.I., MCFHERSON,J., GOLDBERGER,G., KARATHANASIS,S.K. and
 BRESLOW,J.L.
 Synthesis, intracellular processing, and signal peptide of human
 apolioprotein E

JOURNAL J. Biol. Chem. 259 (9), 5495-5499 (1984)
 MEDLINE 84185684
 PUBMED 6325438
 REFERENCE 4 (bases 88 to 1156)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 REFERENCE 5 (bases 577 to 624)
 AUTHORS Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.
 TITLE Isolation and characterisation of a variant allele of the gene for human apolipoprotein E
 JOURNAL Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)
 MEDLINE 85279526
 PUBMED 2992507
 REFERENCE 6 (sites)
 AUTHORS Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.
 TITLE Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant
 J. Clin. Invest. 83 (4), 1095-1101 (1989)
 JOURNAL 89198059
 MEDLINE 2539388
 PUBMED
 COMMENT [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant.
 [5] epsilon-2 allele.
 [6] sites; mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4, epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3],[5]. The sequence shown is 578 homologous with human apo A-I and 818 homologous with rat apo E. For the epsilon-4 sequence, see the separate entry.
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the
 revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)]
 and [3].
 Apo E is located on chromosome 19 --Jackson, Bruns and Breslow , PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1],[J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA.
 Location/Qualifiers
 1. 1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /tissue-type="liver and blood"
 1. 1156
 /gene="APOE"
 /gene="APOE"
 61. 1014
 /codon_start=1
 /product="preapolipoprotein E"
 /protein_id="AAB59546.1"

/db_xref="GI:178851"
 /db_xref="CDB:G00-119-691"
 /translation="MKVLMAALLVFLACQAKVEQAVETEPPELRROOTEMSGORM ELAIGFPMDFYVOTLSROVOEELSSQVQDELRLAMDETKEIKAYSELEDLTP VAEEFRARISKELOAAQARLGDMDVEYCRRLVOYRGEVQAMLGOSTEELRVLASILR KLRRLRLDADDLQKRLAYQAGARGARGLSAIREIRGLPLVEGGRVRAVGSILAG QPIQERNAQMGRLRLARMEMGSRFRRLDEYKEDVAAEVRAKLEEDQAOIIRLOAEAFQ ARLSWFEPLVEDMORWAGLVEKVOAAVGTSAAPVPSDNH"
 61. 114
 /gene="APOE"
 /note="apolipoprotein E"
 115. 1011
 /gene="APOE"
 /product="apolipoprotein E mature peptide"
 267
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
 /citation=[2]
 /replace="c"
 279
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"
 288
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"
 /citation=[4]
 /replace="a"
 369
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"
 409
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="a"
 448
 /gene="APOE"
 /note="in type III hyperlipoproteinemia"
 /replace="c"
 538
 /gene="APOE"
 /note="in type III hyperlipoproteinemia"
 /replace="t"
 568
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="c"
 586
 /gene="APOE"
 /note="in epsilon-2 allele"
 /citation=[3]
 /citation=[5]
 /replace="t"
 702
 /gene="APOE"
 /note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
 /citation=[2]
 /replace="g"
 783
 /gene="APOE"
 /note="in variant"
 /citation=[4]
 /replace="t"
 858
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

Alignment Scores:

Pred. No.: 2.09e-77 Length: 1186
Score: 1573.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.49% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-14 (1-317) x BC003557 (1-1186)

QY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGGTTCTGTGGCTGTGGCTGTGCATCTCCAGAGATGCCAGGCAAGGTG 105
QY 21 GUGlAlaValAlaGlnThrgluProgluProgluLeuArgGlnGlnThrgluTrpGlnSer 40
DB 106 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAGC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GGCACGCGCTGGGAACTGGGCACTGGGCTTTGGGATTACCTGCGCTGGTGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrglnGlnLeuArgAla 80
DB 226 CTGTCTGAGAGGTGGAGAGAGCTGTCTGCTCCAGGTCCAGGAACTGAGGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 286 CTGATGGAGAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACTGGAGGAACAAGT 345
QY 101 ThrProValAlaGlnGlnThrgluThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 346 ACCCGCGTGGCGAGAGAGACGCGGCGACGCTGTCCAGAGAGCTGCAGCGCGCAGGCG 405
QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnThrgluArgGlyGln 140
DB 406 CGGCTGGCGGAGACATGGAGAGCTGTGTGGCGCTGTGTGAGTACCGGCGAGGTG 465
QY 141 GlnAlaMetLeuGlnGlnSerThrglnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCATGCTGCGGAGAGACAGGAGAGCTGCGGCTGCGCTCCCTCCACCTGCGCG 525
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AAGCTCGTAAAGGCGCTCTCCGCGATGCCATGACCTCAGAAAGCGCTGGCAGAGTAC 585
QY 181 GlnAlaGlyAlaArgGlnGlnGlnArgGlyLeuSerAlaHisLeuArgGlnArgGlyGln 200
DB 586 CAGGCGGCGGCGGAGAGGCGCGGCGGCTCAGGCGCATCCGCGAGCGCGCTGCGG 645
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 646 CCCCTGTGGAAACAGGCGCGCTGCGGCGCGCCACATGTGGGCTCCTGGCGCGCGCAG 705
QY 221 LeuGlnGlnArgAlaGlnAlaThrPglLysGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 706 CTACAGAGACGCGGCGCGCTGTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 766 AGCGGAGACCGCGAGCGCTGTGAGAGGTGAAGAGACAGTGGCGGAGTGGCGCGCAG 825
QY 261 LeuGlnGlnGlnAlaGlnGlnAlaArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 826 CTGGAGAGACAGGCGCGAGATAGCGCTGCGAGCGCGAGGCTTCCAGGCGCGCGCTCAAG 885
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlnLysValGlnLys 300
DB 886 ACTGTGTTGAGAGCGCGCTGTGGAGACATGACAGCGCACTGGGCGCGGTGGTGGAGAG 945
QY 301 ValGlnAlaAlaValGlnThrgluSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAAGGCTGCGTGGGACAGCGCGCGCTGTGCGCCAGCGACATATCAC 996

RESULT 8
E08423 1110 bp RNA linear PAT 29-SEP-1997
LOCUS E08423 DNA coding human apolipoprotein E3.
DEFINITION E08423
ACCESSION E08423
VERSION E08423.1 GI:2176540
KEYWORDS JP 1994315392-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1110)
AUTHORS Morimoto,H. and Teranishi,Y.
TITLE METHOD FOR PRODUCING APOLIPOPROTEIN
JOURNAL Patent: JP 1994315392-A 1 15-NOV-1994;
MITSUBISHI KASEI CORP
COMMENT OS Homo sapiens (human)
PN JP 1994315392-A/1
PD 15-NOV-1994
PE 11-JUN-1985 JP 1994015433
PI MORIMOTO HIROMORI, TERANISHI YUNAKA
PC C12P21/02,C07K13/00,C12N5/10,C12N15/12,C12N15/18,C12P21/02,
PC C12R1:91',
CC (C12N5/10,C12R1:91):
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1110
FT sig_peptide 15..68
FT mat_peptide 69..968
FT /product='human apolipoprotein E3'.
FEATURES
source 1..1110
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 198 a 355 c 414 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 4.67e-77 Length: 1110
Score: 1566.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.05% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-14 (1-317) x E08423 (1-1110)
QY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGTTCGTGTGGCTGTGGCTGTGCATCTCCAGAGATGCCAGGCAAGGTG 74
QY 21 GUGlAlaValAlaGlnThrgluProgluProgluLeuArgGlnGlnThrgluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGACAGCAGTGGCAGAGC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACGCGTGGGAACTGGGCACTGGGCTTTGGGATTACCTGCGCTGGTGCAGACA 194
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrglnGlnLeuArgAla 80
DB 195 CTGTCTGAGACAGTGGAGAGAGCTGTCTGCTCCAGTCCAGGAACTGAGGCG 254
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 255 CTGATGGAGAGACCATGAAGAGTTGAAGGCTTCAAAATCGGAATCGAGAACACTG 314
QY 101 ThrProValAlaGlnGlnThrgluThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120

BASE COUNT	208 a	367 c	432 g	149 t
ORIGIN	/db_xref="taxon:3606"			
Alignment Scores:				
Pred. No.:	5,52e-77	Length:	1156	
Score:	1565.00	Matches:	315	
Percent Similarity:	99.37%	Conservative:	0	
Best Local Similarity:	99.37%	Mismatches:	2	
Query Match:	98.99%	Indels:	0	
DB:	6	Gaps:	0	
US-09-827-854-14 (1-317) x BD004277 (1-1156)				
QY	1	McttysvaileutrpalalaleuenuvalThrPhelenualaglCysglnalalasyal	20	
DB	61	ATGAAGGTTCTTGSGGCTGCGTTCGTGTGTCACATTCTCGCAGATGCCAGGCCAAGGTG	12	
QY	21	gluglnalalavalglutThrleuprogluProluenuarGlnglnThrleuTrpGlnSer	40	
DB	121	GAGCAAGCGGTGGAGACAGAGCGCGAGCCGAGCTGCGCACACAGACGAGAGTGCAGAGC	18	
QY	41	glyglinatrTrpGluenualaleuenuLysarPheTrpAspTrpLysarGlyIleThr	60	
DB	181	GGCCAGCGGTGGAACTGGCACTGGGTGGCTTTTGGAATTACTCTCGCTGGGTGGACACA	24	
QY	61	LeuSerGluGluIValGlnglnGluLeuLeuSerSerGlnIValThrGlngluenuarGla	80	
DB	241	CTGTCTGACACAGGTGCAGAGAGAGAGCTGCTCACTCCACAGTACCAGAACTAGAGGCG	30	
QY	81	IeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnglnleu	100	
DB	301	CTGATGGAGACAGACCATGAGAGAGATTGAAGGCTTCACAATCGAAGCTGAGAACACTG	36	
QY	101	ThrProValAlaGluGluThrArgAlaArGleuSerLysGluLeuGlnAlaAlaGlnAla	120	
DB	361	ACCCCGGTGGCGGAGAGACGCGGGCACGGCTGTCTCAAGGACTCCAGGCGCGCAGGCGC	42	
QY	121	ArgLeuGlyAlaAspMetGluAspValArgGlyArGleuValGlnTrpArgGlyGluVal	140	
DB	421	CGGCTGGGCGCGACATGGAGAGACTGTGCGGCGCCCTGTGTGCACTACCCGCGCAGGTG	48	
QY	141	GlnAlaMetLeuGlyGlnSerThrGlnglnGluLeuArGlaArgLeuAlaSerHisLeuArg	160	
DB	481	CAGGCCATGCTCGGCGCAGAGCACGAGAGAGTGGGGGTGGCGCTCGCCACCATCGGCG	54	
QY	161	LysLeuArGlyArGleuLeuArGlyAspAlaAspAspLeuGlnLysArGleuAlaValTrp	180	
DB	541	AAGCTGCGTAAACGGGCTCTCCGCAATGCGCATGACCTGCACAACTGCTGGCAATTGAC	60	
QY	181	GlnAlaGlyAlaArgGlnglyAlaGluArgGlyLeuSerAlaIleArgGluArGleuGly	200	
DB	601	CAGGCCGGGGGCCCGGAGAGGCGCCGACGGCGGCTTCAGAGGCCCATCCGCGAGCGCTGGGG	66	
QY	201	ProLeuValGluGlnGlyArGlyArGlyAlaArGlyAlaTrpValGlySerLeuAlaGlyInPro	220	
DB	661	CCCGTGGTGGAGACAGGGCGCGCTGGGGGCGCCACTGTGGGGTCCCTGGCGGCGCAGGCG	72	
QY	221	LeuGlnGluArGlaGlnAlaTrpGlyLeuArGleuArGlaArGleuMetGluMetGly	240	
DB	721	CTACACGAGAGCGGCGCCAGGCTCTGGCGCAGCGGCTGCGCGCGGATGAGAGATGGGCG	78	
QY	241	SerArGlyThrArgAspArLeuAspGluValLysGlnglnIValAlaGluValArgAlaLys	260	
DB	781	ACCAGGACCCGCGACCGCTTGACACAGGTAAGAGAGACAGGTGGCGGAGGTGGCGCCACAG	84	
QY	261	LeuGlnGluGlnAlaGlnIleArGleuGlnAlaGlnAlaPheGlnAlaArGleuLys	280	
DB	841	CTGGAGAGAGACAGCCACAGATACGCTGCAGCGCGCAGGCGCTTCAGAGCCGCGCTCAAG	90	
QY	281	SerTrpPheGluTrpLeuValGluAspMetGlnArGlnTrpPalaGlyLeuValGluLys	300	
DB	901	ACCTGTTTCGAAACCCCTGGTGGAAACAATGCAAGCCCAAGTGGGCGCGGCTGTGTGAAGAG	96	

QY 301 ValGlnAlaAlaValAlGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGTGGGACACGAGCGCGCCGTGTGCCAGCAGCATCTAC 1011
RESULT 10
AX333278
LOCUS AX333278 1157 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 3787 from Patent WO0194629.
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigian, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 6.26e-77 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 3
Query Match: 98.92% Indels: 0
Gaps: 0
US-09-827-854-14 (1-317) x AX333278 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATAAAGGTTCTGTGGCTGCTGTGCTGCTACATCTCTGGCAGATGCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValAlGlyThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPglIns 40
Db 122 GACCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCCAGAGACCGAGTGCAGAGC 161
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGGCAGCGCTGGGAGACTGGCTGCTGCTTTGGGATTACCTGGCTGGTGCAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGCAAGAGAGCTGCTCAGCTCCCAAGTCCCAAGAGACTGAGGGCG 301
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGAGAGAGACATGAAGAGCTTGAAGGCTTACAAATCGAAGTGGAGGAACAAGT 361
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGAGAGCGCGGCAAGCTGTCCCAAGAGCTCCAGACGCGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnThrArgGlyGlnVal 140
Db 422 CGGCTGGGCGGAGACATGAGAGAGTGTGGCGCGCTGTGACATGACGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnLysInsThrGlnGlnLeuArgValArgPheAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGCGGACAGAGACCGAGAGCTCGGGTGGCTCCCTCCCACTGGCGG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

Db 542 AACCTGGTAAAGCGGCTCCCTCCGATCCGATGACCTCGAAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 602 CAGGCGGCGGCGCGGAGAGCGCGGAGCGGCGGCTCAGGCGCATCCCGAGCGCGCTGGG 661
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
Db 662 CCCCTGGTGAACAGGCGCGGCTGGCGGCGCGGAGCTGTGGCTGCGCGCGGCGAGCGG 721
QY 221 LeuGlnGlnArgAlaGlnAlaThrPglGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 722 CTACAGAGCGGCGCGGAGCGGCTGGCGGAGCGGCTGGCGGAGTGAAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAla 260
Db 782 AGTCGAGACCGGAGCGCGGCTGGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
Db 842 CTGAG 901
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 902 AGCTGTTGAGGCGCGGCTGGTGAAGACATGACAGCGGAGGCGGCGGCTGTGTGAAGAG 961
QY 301 ValGlnAlaAlaValAlGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGTGGGACACGAGCGCGCCCTGTGTGCCAGCAGCATCTAC 1012
RESULT 11
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS AX409597
DEFINITION Sequence 2244 from Patent WO0229103.
ACCESSION AX409597
VERSION AX409597.1 GI:21442302
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2244 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M12529"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 6.26e-77 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 3
Query Match: 98.92% Indels: 0
Gaps: 0
US-09-827-854-14 (1-317) x AX409597 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATAAAGGTTCTGTGGCTGCTGTGCTGCTACATCTCTGGCAGATGCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValAlGlyThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPglIns 40
Db 122 GACCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCCAGAGACCGAGTGCAGAGC 181

Oy		41	Glyglnatrgtppglulealaleuengllyarphetraparyrleuargrtryala1nthr	60
Db		182	GGCCAGCCCTGGGAACGTGGCACTGGGTGCTTTTGGATTAACAATCGCTGGGTGCAGACA	241
Oy		61	LeuSerGIugInValGlnglunleuLeuSerSergInValThnglunleuArrgla	80
Db		242	CTGTGTGAGCAGAGTCACAGAGAGCTGCTCAACTGCCAAGTACCACAAAGAACAGAGGGCG	301
Oy		81	LeuMetaspGIurthMetIysgluleuLysalarTrlysSserGIuLeungluInleu	100
Db		302	CTGATGGACGACACCATGAGAGATTGTAAGGCGCTCAAAATCGGAATCGAGAACAACTGG	361
Oy		101	ThrProVALAlagluIngluthArghlarArgleuSerIySGluLeuGlnAlalaelnla	120
Db		362	AACCCGGTAGGGGAGGAGAACGGGGCACGGCTGTCCAAGAGACTCAACAGCGGAGGCC	421
Oy		121	ArgleuGIyAlaAspMetGIunAspyAlArgeIyArgleuValGIntYrArgIygluVal	140
Db		422	CGGCTGGGGCGCGAATGAGAGACTGTGCGGGCCGCTGGTGCAATACCAGCGGCAGGTG	481
Oy		141	GlnAlaMetLeuGIyGlnSerThrGIuInleuArvalArgleuAlaSerHisLeuArg	160
Db		482	CAGGCCATGCTCGGCAGACAGCACAGAGAGCTGGGGGTGGCGCTGCCCTCCACCTGGCG	541
Oy		161	LysLeuArGIySArygleuLeuArqspAlaspsApbleuGlnLysArgleuAlaValIyr	180
Db		542	AAGCGCTTAAGCGGGCTCTCCGCAATCCCGATACCTGCAGAAACGGCTGGCACTGTAC	601
Oy		181	GlnAlaGIyAlaArqGIuGIyAlaGIuArgIyArgeIyLeuSerAlaIlleArGIuArgleuGIy	200
Db		602	CAGGCCGGGGCCCCGAGAGGGCCCGAAGCGCGGCTCAAGCGCCATCCGAGAGCGCTGGGG	661
Oy		201	ProLeuValGIuIngInGIyArqValArqlaalathrValGIySerIeuaAlaglInPro	220
Db		662	CCCCTGTGTGAACAGGGCGCGGTGGGGCGCCACTGTGGGTCTCCCTGGCGGGCAGCGG	721
Oy		221	LeuGIuIngluArqlaGIuAlaAtrrpelyluargleuArqAlaArgMetGIuImetGIy	240
Db		722	CTACAGGAGCGGGCCCAAGGCTGTGGGGCCAGCGGCTGCAGCGCGGATGGAGAGATGGGC	781
Oy		241	SerArqTrArqAspArgIeuaSpdluValIySGluIngluInValAlaGIuValArqlAlays	260
Db		782	AGTCGGACCCCGGACCGCTGTGGACGAGGTGAAGAGCAAGGTGGCGAGGTGGCCGCCXAG	841
Oy		261	LeuGIuIngluInlaIngInIlleArqLeuGlnAlaGlnAlaArheGlnAlaArgLeuIys	280
Db		842	CTGGAGGAGCAGGGCCACAGATACGGCTGCAGGGCGAGGGCTTCACAGGCCGCTCAAG	901
Oy		281	SerTrpHeuIuProlEuValGIuAspMetGlnArqGIntrrpAlaGIyLeuValGIuIys	300
Db		902	AGCTGGTTCGAGCCCTGTGTGAANAACATGTGACAGCGCCAGTGGGGCGGCTGTGTGAGAAG	961
Oy		301	ValGlnAlaAlaValGIyThrSerAlaIaIaProlValProSerAspAsnHis	317
Db		962	GTGGAGGCTGCGGTGGGCACAGCGCCGCCCTGTGTGCCAGGACAAATCAC	1012
RESULT	12			
LOCUS	IL15975	115975	1157 bp	DNA linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.			
ACCESSION	U15975			
VERSION	U15975.1	GI:1250883		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1157)			
TITLE	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.			
JOURNAL	Production of recombinant proteins in insect larvae			
FEATURES	Patent: US 5472858-A 1 05-Dec-1995;			
	Location/Qualifiers			
	1..1157 /organism="unknown"			

BASE COUNT	212 a	370 c	426 g	149 t	ORIGIN
Alignment Scores:					
Pred. No.:	6,26e-77			1,157	
Score:	1564.00			314	
Percent Similarity:	99.05%			conservative: 0	
Best Local Similarity:	99.05%			mismatches: 3	
Query Match:	98.92%			gaps: 0	
DB:	6			0	
US-09-827-854-14 (1-317) x IL5975 (1-1157)					
QY	1	MctlysvaLeutrrpaAlaLeuLeuValthrPheLeuAlaGlyCysGlnAlaLysVal	20		
DB	62	ATGAAGGTTCTGTGGGCTGGCTGTGTATTCATTCTCGAGGATCCAGGCCAAGGTG	121		
QY	21	GtGlnAlaValaGltThgGlpProGluProGluLeuArgGlnGlnTrpGlnTrpGlnSer	40		
DB	122	GAGCAACGGCGTGGACACAGAGCGGAGCGCCAGCGCCGACAGACAGCGAATGGCACAGC	181		
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyLyrPheTrpAspTyrLeuArgTrpValGlnThr	60		
DB	182	GGCCAGCGCTGGGAAGCTGGCACTGGGTGGCTTTTGGATTCACGCGCTGGGTGCACACA	241		
QY	61	LeuSerGluGlnValGlnGlnGluLeuLeuSerGlnValTrpGlnGluLeuArgAla	80		
DB	242	CTGTCTGACAGAGTCCAGAGAGAGCTGTCAAGTCCCAAGTCACCCAAAGACTAGGGCG	301		
QY	81	LeuMetAspGluThrMetLeuGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu	100		
DB	302	CGATGGAGAGACCATGAGAGAGATTGAAGCCCTTCAACAATCGGAATCGAGAACACTG	361		
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
DB	362	ACCCCGGTAGCGCGAAGACACGGGGCGACGGCTGTCTCAAGGAGCTGCAGACGGCGCAGGCC	421		
QY	121	ArgLeuGlyAlaAspMetGluAspValAlaArgGlyLyrGluLeuValGlnTyrArgGlyGluVal	140		
DB	422	CGGCTGGGCGCGACATGAGAGAGCTGTGGCGCGCTGTGTCAATCCCGCGGAGGTG	481		
QY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuValArgValAlaGluLeuAlaSerHisLeuArg	160		
DB	482	CAGGCCATCTCGGCGACAGACACCGAGAGAGCTGGGGTGGCGCTCCGCTCCACCTCGGC	541		
QY	161	LysLeuAlaGlyLysArgTrpLeuAlaArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180		
DB	542	AAGCTGCGTAAACGGCTCTCTCCGCATCCCATGACCTGCAGAAAGCGCTGGCAGGTATC	601		
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluGluArgGlyLeuSerAlaAlaLeuArgGluLeuGly	200		
DB	602	CAGGCCGGGGCGCGGAGAGGCGCCGAGCGCGGCTCAAGCGCATCCCGGAGCGCTGGGG	661		
QY	201	ProLeuValGluGlnGlnLyrGluValAlaAlaThrValGlySerLeuAlaGlyLysPro	220		
DB	662	CCCTGGTGGAAACAGAGGCGCGGTGGCGGCCGCCCACTGGGCTCCCTGGCGGCCACAGCG	721		
QY	221	LeuGlnGluAlaArgAlaGlnAlaAlaTrpGlyLysLysArgLeuArgAlaAlaGlyMetGluGly	240		
DB	722	CTAACAGAGAGCGGCGCCAGCGCTGGGGCGACGGCTGGCGCGCGCGAGTGGAGACATGGGC	781		
QY	241	SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys	260		
DB	782	AGTCGAGACCCCGAGACCGCTTGACAGAGTGAAGAGCAAGTGGCGGAGGTGCCGCCAAG	841		
QY	261	LeuGlnGluGlnAlaGlnGlnLysLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280		
DB	842	CTGGAGGAGACAGGCCACACAGATACGCTTCGACGCCAGAGGCTCCAGGCCCGCTCAAG	901		
QY	281	SerTrpPheGluLyrProLeuValGluLysPheClnArgGlnTrpAlaGlyLeuValGluLys	300		
DB	902	AGCTGGTTCGAGCCCTGTGTGGAAGACTGCAGGCGCAAGTGGCGCGGGCTGTGTGAAG	961		

QY	301	VALGINAIAAValGlyThrsSerLaaIAPcValProSerAspAsnHis	317
Db	962	GTGCAGGCTGGCGGCACAGCGCCGCCCTGTGCACGACAAATCATC	1012
RESULT	13		
LOCUS	HUMAPOE	1157 bp	MRNA
DEFINITION	Human apolipoprotein E mRNA, complete cds.		linear
ACCESSION	M12529		PRI 08-AUG-1995
VERSION	M12529.1	GI:178848	
KEYWORDS	apolipoprotein.		
SOURCE	Homo sapiens (clone: PHAE[112,178,813].)	male 57-year old liver	
ORGANISM	CDNA to mRNA.		
REFERENCE	Homo sapiens		
AUTHORS	McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.		
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant		
JOURNAL	J. Biol. Chem.	259 (10), 6498-6504	(1984)
MEDLINE	84212473		
PUBMED	6327682		
FEATURES			
Source	Location/Qualifiers		
	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="19q13.2"		
	/clone="PHAE[112,178,813]."		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="57-year old"		
	1..1157		
	/gene="APOE"		
	<1..1157		
	/gene="APOE"		
	/note="apoE mRNA; G00-119-691"		
	62..1015		
	/gene="APOE"		
	/note="precursor"		
	/codon_start=1		
	/product="apolipoprotein E"		
	/protein_id="AAB59518.1"		
	/db_xref="GI:178849"		
	/db_xref="GDB:G00-119-691"		
	/translation="MKVMAALVTPLPAGCOAKVEQAVETPEPELROQTFMWSGGQMM ELAGRGWDYLRMYOTLSSEQVQBELSSQVQELRALMDERMKELKAKSLSEQLP VAETRRRLSKELDTAQRRLDADMEDVCGRLVQTRGEVQALGQSTELRVLRASHLR KLRLRLNDPDDLOKRLAVYQAGAEQAEGLSAIRRLGLVQGRVATVYSLAG QPLDERQAWGERLRARMEGSRTRRLDEKRVQAVEVRAKLEEQAOQIRLQAEARQ ARLSWEFLVEEDMQRWQAGLVEKVAAGTSAAPVPSDNH"		
	62..115		
	/gene="APOE"		
	/note="G00-119-691"		
	116..1012		
	/gene="APOE"		
	/product="apolipoprotein E"		
	/note="G00-119-691"		
	370		
	/gene="APOE"		
	/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	410		
	/gene="APOE"		
	/note="a in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	569		
	/gene="APOE"		
	/note="c in PHAE[112,178]; g in PHAE813; G00-119-691"		
	/replace="g"		
	784		
	/gene="APOE"		

variation	/note="t in PHAE[112,178]; c in PHAE813; G00-119-691" /replace="c" 859 /gene="APOE" /note="g in PHAE[112,178]; a in PHAE813; G00-119-691" /replace="a"
BASE COUNT	212 a 370 c 426 g 149 t
ORIGIN	30 bp upstream of BstNI site.
Alignment Scores:	
Align. No.:	6,26e-77
Score:	1564.00
Percent Similarity:	99.05%
Best Local Similarity:	99.05%
Query Match:	98.92%
DB:	9
US-09-827-854-14 (1-317) x HUMAPOE (1-1157)	
0y	1 MettysValLeutrpAlaAlaLeuLeuValThrpHeuAlaGlyGlnAlaLysVal 20
Db	62 ATGAAGGTTCTGTGGGCTGCTTCTGTGTCATTTCTTGCGAGATGCCAGGCCAAGTGT 121
0y	21 GluGlnAlaValAlaGluThnGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
Db	122 GAGCAAGGGGTGGAGACAGACGCCGAGCCCAAGCTGGCGCCAGCGAGACCGAGTGGCAGAGC 181
0y	41 GlyGlnArGTrpGluLeuAlaLeuGlyArpPhrTrpAspTrpLeuArGTrpValGlnThr 60
Db	182 GGCACGCGCTGGGAAGTGGCACTGGTGTGCTTTGGGATTACCTGGCGCTGGGTGCAGACA 241
0y	61 LeuSerGluGlnValGlnGluLeuLeuLeuSerGlnValThrnGlnGluLeuArgAla 80
Db	242 CTGTCTGACGAGGTGCAGAGAGAGACTGTCTGCTCCCAAGTCACCCAAAGAACTGAGGGCG 301
0y	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
Db	302 CTGATGGACGAGACCATTGAAGAGCTTGAAGCCCTTACAAATCTGGAACCTGGAGAAACAATG 361
0y	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db	362 ACCCGCGTAGCGGAGAGAGACCGGGGCAAGCTGTGCCAAGAGCTGCAGACGCGCGCAGGCC 421
0y	121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140
Db	422 CGGCTGGCGCGGACATGAGAGAGCTGGCGCGCGCTGGCGCAGTACCAGCGGAGAGTGT 481
0y	141 GlnAlaMetLeuGlyGlnSerThnGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db	482 CAGGCAATGCTGGCCAGACAGCACCGAGAGAGCTGGGGTGGCGCTCGCTCCACATCGCC 541
0y	161 LysLeuArGTrpArgLeuLeuArGAspAlaAspAspLeuGlnLysArgLeuAlaValAlaTrp 180
Db	542 AAGCTGCGCTAAGCGGCTCTCTCCGGATCCCGATACCTGCAGAAAGCGCTGGCAGTGTAC 601
0y	181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db	602 CAGGCGGGGGCGCGGAGAGCGCGGAGCGGCTCTCAGCGCATCCCGAGAGCGCTGGGG 661
0y	201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db	662 CCCCTGGTGGAAACAGGGCGGGTGGGGCGCCCATCTGGGCTCTCCCTGGCGGCGCACCG 721
0y	221 LeuGlnGluArGTrpAlaGlnAlaTrpGlyGluLysArgLeuArgAlaArgMetGluGluMetGly 240
Db	722 CTACAGAGCGCGGCGCCAGGCTGGGGCGAGAGCGCTGGCGCGCGCGAGTGAAGATGGCC 781
0y	241 SerArGTrpArGAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db	782 AGTGGAGACCGCGACCGGCTGAGAGGTGAAGAGCAGGATGGCGGAGTGGCGGCAAG 841
0y	261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280

Db 842 CTGGAGACGACCCAGCAGATACGCTTCAGAGCCCTTCAGGCCCTCAAG 901
Oy 281 SerTrrpHegluProleuValGluaspMetGlnArgGlnTrrpAlaGlyLeuValGluys 300
Db 902 AGCTGTTGACGCCCCGTGGAGACATGACGCGCAGTGGCGGCTGGAGAG 961
Oy 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCCGTGGGACACAGCGCGCCCTGTGGCCAGGACATCAC 1012
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS Monkey mRNA for apolipoprotein E.
DEFINITION X13887
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
REFERENCE 2 (bases 1 to 1178)
AUTHORS Marotti K.R., Whitted B.E., Castle C.K., Polites H.G. and
Meichlor G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PubMed 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/cissue_type="liver"
/clone_lib="lambda gt11"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_18="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTVFLAGCOAKVEOPVEPEPELROAGGSGOPM
ELALGRFMDYLRWVOTLSEVOEELISPOVQOELTTLMDETMKELKAYKSLLEOLSP
VAETRRARLSKELOAAOARLADMEDEVASRLVORSEYOAMLGSTELRRLASHLR
KLKRLDLADDDQRLAVYQAGAREGERSVSIKRELGFLVQGRVRAATVGSLS
QPLDERAQLALVEERLARMEMGSRTRDLDEKQVAVERRAKLEEQAOQISLQEARQ
ARLSWEPEPLVEDMQRQWAGLVERKQAAVASTAPVPIDNH"
83..136
sig_peptide 137..1033
mat_peptide /product="apolipoprotein E (AA 1 - 299)"
1154..1159
misc_feature /note="polya signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 3,17e-71 Length: 1178
Score: 1459.00 Matches: 297
Percent Similarity: 94.32% Conservative: 2
Best Local Similarity: 93.69% Mismatches: 18
Query Match: 92.28% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-14 (1-317) x MFAPOE (1-1178)
Oy 1 MetLysValLeuTrrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 83 ATGAAGGTTCTGTGGGCTGCGTGTGTCATCTCGGAGATGCGAGGCAAGGTG 142
Oy 21 GtuglnAlaValAlaGlnThrGluProGluProGluLeuAlaGlnGlnThrGluTrrpLnsr 40
Db 143 GAGCAACCGGTGGAGCCAGAGACAGACCCGAGCTTCGCGACAGCTGAGGCGAGAC 202
Oy 41 GlyGlnArgTrrpGlnLeuAlaLeuGlnLysArgPheTrrpAspTrrpLeuArgTrrpValGlnThr 60
Db 203 GCCAGCCCTGGAGACTGGCACTGGGTGCTTTTGGATTACCTGCGTGGGTGAGACA 262
Oy 61 LeuSerGlnGlnValAlaGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuValArgAla 80
Db 263 CTGTGTAGACAGCTGACAGAGAGAGCTGCTCACCCCCAGAGTCAACCCAGGAAGTACAGAC 322
Oy 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTrrpLysSerGlnLeuGlnGlnLeu 100
Db 323 CTGATGACAGAGACCATGAAAGAGTGAAGGCTTCAAATCGGAACCTGGAGACAGCTG 382
Oy 101 ThrProValAlaGlnGlnThrPrrgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 383 AGCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAAAGAGCTCAGCGCGGCGAGGCC 442
Oy 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnTrrpArgGlyGlnVal 140
Db 443 CGGCTGGGTGCCGACATGAGAGAGCTGCGGACGCCCTGTGTCACTACCGCAGGAGGTG 502
Oy 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGCATGCTGGGCGCAAGTACCAGAGAGCTGCGGCGGCGCTGCCCTCCACCTGGCG 562
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 563 AAGCTGCCAGACCGGCTCTCCGCGATGTGATGACTTGCAGAAAGCGCTTGCACTGTAT 622
Oy 181 GlnAlaGlyAlaArgGlnGlnValAlaGlnLysArgGlyLeuSerAlaAlaArgGlnLysGln 200
Db 623 CAGCGCGGCGGCCCGGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 682
Oy 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 683 CCCCTGGTGGAGACAGGCGCGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
Oy 221 LeuGlnGlnArgAlaGlnAlaTrrpGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 743 CTTCAGGAGCGGCGCCAGGCGCTTGGGTAGCGGCTTCGCGACAGGATGGAGAGATGGGC 802
Oy 241 SerArgTrrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
Db 803 AGCCGAGCCCGGACCGGCTGTGAGCAGGTGAAGAGACAGGTGGCGAGGTGCGGCCCAAG 862
Oy 261 LeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 863 CTGCAGGAACAGCGCCACAGATAGCTGCAGGCGGAGGCGCTTCAGAGCCGCTCAAG 922
Oy 281 SerTrrpHegluProleuValGlnaspMetGlnArgGlnTrrpAlaGlyLeuValGluys 300
Db 923 AGCTGTTGACGCCCCGTGGAGACATGACGCGCAGTGGCGGCTGGAGAG 982
Oy 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGCAGGCTGCCGTGGGACACAGCGCGCCCTGTGGCCAGGACATCAC 1033
RESULT 15
HUMAPOE4 5515 bp DNA linear PRI 09-NOV-1994
LOCUS Human apolipoprotein E (epsilon-4 allele) gene, complete cds.
DEFINITION M10065 J030053 J030054
ACCESSION M10065.1 GI:1178852
VERSION M10065.1 GI:1178852
KEYWORDS Alt repeat; allelic variation; apolipoprotein; apolipoprotein E;
lipoprotein; repeat region; very low density lipoprotein.
SOURCE Human DNA [2], [1].
ORGANISM Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 5515)				
AUTHORS	Das,H.K., McPherson,J., Bruns,G.A., Karathanasis,S.K. and Breslow,J.L.				
TITLE	Isolation, characterization, and mapping to chromosome 19 of the human apolipoprotein E gene				
JOURNAL	J. Biol. Chem. 260 (10), 6240-6247 (1985)				
MEDLINE	85207610				
PUBMED	3922972				
REFERENCE	2 (bases 196 to 5269)				
AUTHORS	Paik,Y.K., Chang,D.-J., Reardon,C.A., Davies,G.E., Mahley,R.W. and Taylor,J.M.				
TITLE	Nucleotide sequence and structure of the human apolipoprotein E gene				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 82 (10), 3445-3449 (1985)				
MEDLINE	85216517				
PUBMED	2987927				
REFERENCE	3 (bases 1 to 5515)				
AUTHORS	Emil,M., Wu,L.L., Robertson,M.A., Myers,R.L., Hegele,R.A., Williams,R.R., White,R. and Lalouel,J.M.				
TITLE	Genotyping and sequence analysis of apolipoprotein E isoforms				
JOURNAL	Genomics 3 (4), 373-379 (1988)				
MEDLINE	89212602				
PUBMED	3243553				
COMMENT	[3] two allelic variations. Draft entry and computer-readable sequence for [3] kindly provided by M.Emil, 19-AUG-1988.				
FEATURES	Apolipoprotein E is a constituent of the human very low density lipoprotein in the plasma. There are at least six distinct phenotypes derived from the single E gene on chromosome 19; next to the epsilon-3 allele (see separate entry), the epsilon-4 allele, represented by the sequence below, is most common, the product difference being arginine in place of cysteine at residue 112 [2]. The gene structure of apo E is similar to that of other apo genes: presence of the 66-bp repeats in the fourth exon (starting at base 3782 below) makes the E gene highly similar to the A-I gene (see separate entry) as argued by [1].				
FEATURES	A potential TATA box is found at positions 1014-1018, and a potential polyadenylation signal at 4616-4621.				
FEATURES	[2] and [1] had slight differences in the boundary positions for the Alu repeats and their flanks: the boundary positions indicated in [1] have been used in the FEATURES table below. Draft entries and clean copies were kindly supplied by J.M. Taylor, Gladstone Laboratories, San Francisco, and by J.P. Levine, Rockefeller University, New York.				
FEATURES	Location/Qualifiers				
source	1..5515				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
conflict	/map="19q13.2"				
	202..203				
	/citation=[2]				
	/replace=""				
repeat_region	complement(300..345)				
	/note="direct repeat flanking Alu repeat 3' copy [1]"				
repeat_region	complement(346..635)				
	/note="Alu repeat [1]"				
repeat_region	complement(636..680)				
	/note="direct repeat flanking Alu repeat 5' copy [1]"				
conflict	963..967				
	/citation=[2]				
	/replace=""				
exon	1047..1090				
	/note="apo E mRNA [2],[1]"				
	/number=1				
intron	1091..1847				
	/note="apo E mRNA intron 1 [2],[1]"				
conflict	1362..1363				
	/citation=[2]				
	/replace=""				
conflict	1789..1793				
	/citation=[2]				
CDS	/replace=""				
	join(1871..1913,3007..3199,3781..4498)				
	/note="precursor"				
	/codon_start=1				
	/product="apolipoprotein E"				
	/protein_id="AA059397.1"				
	/db_xref="GI:178853"				
	/translation="MKVLAALLVFLLACCAKVEQAVETPEPELROQTEWQSGCRM ELALGFWDYLRWVOTLSEQVDELSSQVTOELALMDETKELKAVKSELEOLTP VAETPARLSKETLAAQARLGGADMEDVRSGLVYRGEOAMLGOSTEELRVLASHLR KLRRLRDADLQRLAYVQAGAREGASAIERLGPLVEQGRVAAYVGSILAG OPLQEOAQCERLRAEMEGSRTPEDRLDEYAEVRAKLEEDAAQIRLQAEAFQ ARKSWFEPLVEDMQRNAGLVEKVOAAGTSAAPVSDNH"				
	join(1871..1913,3007..3017)				
	join(1871..1913,3007..3781,4495)				
gene	/product="apolipoprotein E"				
	1871..1913				
	/gene="APOE"				
exon	1871..1913				
	/partial				
	/gene="APOE"				
	/note="preapolipoprotein E; G00-119-691"				
	/number=2				
intron	1914..3006				
	/note="apo E cds intron 2 [2],[1]"				
repeat_region	complement(2092..2104)				
	/note="direct repeat flanking Alu repeat 3' copy [1]"				
repeat_region	complement(2105..2429)				
	/note="Alu repeat [1]"				
repeat_region	complement(2430..2442)				
	/note="direct repeat flanking Alu repeat 5' copy [1]"				
repeat_region	2520..2526				
	/note="direct repeat flanking Alu repeat 5' copy [1]"				
repeat_region	2527..2886				
	/note="Alu repeat [1]"				
conflict	2568..2569				
	/citation=[2]				
	/replace=""				
repeat_region	2887..2893				
	/note="direct repeat flanking Alu repeat 3' copy [1]"				
conflict	2947..2948				
	/citation=[2]				
	/replace=""				
conflict	2983..2984				
	/citation=[2]				
	/replace=""				
exon	3007..3199				
	/number=2				
variation	3182				
	/note="g in epsilon-4; a in epsilon-3 [2]"				
variation	3191				
	/note="g in epsilon-4; a in epsilon-3 [2]"				
intron	3200..3780				
	/note="apo E cds intron 3 [2],[1]"				
conflict	3568..3569				
	/citation=[2]				
	/replace=""				
exon	3781..4640				
	/partial				
	/note="preapolipoprotein E"				
	/number=3				
variation	3817..3819				
	/note="cta in [2] and 1 allele [3]; agc in other allele [3]"				
variation	3844..3846				
	/note="gac in [2] and 1 allele [3]; tgt in other allele [3]"				
variation	3932				
	/note="c in epsilon-4; t in epsilon-3 [2] (arg; cys)"				
variation	4342				
	/note="g in epsilon-4; a in epsilon-3 [2]"				
conflict	4708..4709				
	/citation=[2]				

```
/replace=""
repeat_region complement(4761..4768)
/replace="direct repeat flanking Alu repeat 3' copy [1]"
repeat_region complement(4769..5048)
/replace="Alu repeat [1]"
repeat_region complement(5049..5056)
/replace="direct repeat flanking Alu repeat 5' copy [1]"
BASE COUNT 1042 a 1667 c 1600 g 1206 t
ORIGIN 201 bp upstream of BanII site on chromosome 19q12-q13.2.

Alignment Scores:
Pred. No.: 1.24e-66 Length: 5515
Score: 1387.00 Matches: 307
Percent Similarity: 58.22% Conservative: 1
Best Local Similarity: 58.03% Mismatches: 8
Query Match: 87.73% Indels: 214
DB: Gaps: 2

US-09-827-854-14 (1-317) x HUMAPOB4 (1-5515)
Qy 2 LysValLeuTrpAlaAlaLeuVal----- 10
Db 2910 AAGCATTGTGTGGAGACCTTCTGTGTGCTAGCTAGCTAGATGCTGGACGGGGCTAG 2969
Qy 11 -----ThrpheLeuAlaGlyCysGlnAlaLysValGln 22
Db 2970 AAGGACCTGAGCCGACCTTGACTGTTCACACAGAGATGCCAGGCAAGGTGAGCA 3029
Qy 23 AlaValGluThrGluProGluProGluLeuArgGlnThrGluThrGlnSerGlyGln 42
Db 3020 GCGGTGGAGACAGAGCCGAGCCGAGCTGCGCGCAGCAGACGAGTGGCAGAGCGGCAG 3089
Qy 43 ArgTrpGluLeuAlaLeuGlyArgThrPaspTrpLeuArgTrpValGlnThrLeuSer 62
Db 3090 GCGGTGGAGACGAGCTGCGGTGCTTTGGGATTTACCTGCGCTGGCTGACAGACTGTCT 3149
Qy 63 GluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeu----- 78
Db 3150 GAGCAGGTGACAGAGAGACTGCTGACGTCACAGTCCAGCCAGAACT-GAGGTAGTGTCTC 3208
Qy 78 ----- 78
Db 3209 CCCATCTCGCCCTTGACCTCTGCTGGGGGCTATACCTCCAGGTCAGGTTTCAT 3268
Qy 78 ----- 78
Db 3269 TCTGCCCCCTTGCGCTAAGTCTTGGGGGCGCTGGGTCTCTGCTGTTAGCTTCTCTTC 3328
Qy 78 ----- 78
Db 3329 CCATTTCGACTCGCTGCTTAAAGTCTGTGAATTCTCTCTGACGTTTGTCTCTCTCT 3388
Qy 78 ----- 78
Db 3389 CTTCCTTCTGACTAGTCTCTACACTGCTGCTGCTGTCTGTCTGTCTCTCTCTCTCT 3448
Qy 78 ----- 78
Db 3449 CTTTATATAGAGACAGAGAGATGGGGTCTCACTGTGTGCCCAGGCTGGTCTTGAACCTT 3508
Qy 78 ----- 78
Db 3509 CTGGGCTCAGAGAGATCTCTCCGCGCTCGCCCAAGTGTGGGATTAGAGCATGAGC 3568
Qy 78 ----- 78
Db 3569 ACCTTGCCGGGCTCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3628
Qy 78 ----- 78
Db 3629 TGCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3688
Qy 78 ----- 78
```

```
Db 3689 CTCCTGCGTCATCCCATCTCGCCGCCCATCCAGCCCTTCTCCCCCGCTCCAC 3748
Qy 79 -----ArgAlaLeuMetAspGlnThrMetLysGlu 88
Db 3749 TGTGGACACCCCTCCGCCCTCTCGCGCCAGGGCGGTAGTGAAGACCATGTAAGAG 3808
Qy 89 LeuLysAlaTrpLysSerGluLeuGluGluLeuThrProValAlaGluThrArg 108
Db 3809 TTGAAGGCTTACAAATCGGAACGTGAGAGACACTGACCCCGGTGGCGAGAGCGCG 3868
Qy 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGluAlaAspMetGluAsp 128
Db 3869 GCACGCGTGTCCAAAGAGACTGACAGCGCGCAGCGCCCGGTGGCGCGCATGAGAGAC 3928
Qy 129 ValArgGlyArgLeuValGlnThrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 3929 GTGCGCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3988
Qy 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysLeuArg 168
Db 3989 GAGGAGCTGCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4048
Qy 169 AspaLysAspLeuGlnLysArgLeuAlaValTrpGlnAlaGlyAlaArgGluGlnVal 188
Db 4049 GATGCCGATGACCTGACAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4108
Qy 189 GluArgGlyLeuSerAlaIleArgGluArgLeuGlyProLeuValGluGlnGlyArgVal 208
Db 4109 GAGCGCGCGCTGACGCGCATCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCTG 4168
Qy 209 ArgAlaIleThrValGlySerLeuAlaGlyGlnProLeuGlnArgAlaGlnAlaTrp 228
Db 4169 GGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4228
Qy 229 GlyGluArgLeuArgAlaArgMetGluGluMetGlySerArgTrpArgAspArgLeuAsp 248
Db 4229 GCGCAGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 4288
Qy 249 GluValLysGluGlnValAlaGluValArgAlaLysLeuGluGluGlnAlaGlnIle 268
Db 4289 GAGGTGAAGAGAGAGTGGCGGAGGTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 4348
Qy 269 ArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLysSerTrpPheGluProLeuValGlu 288
Db 4349 GCGCTGCGAGCGCGAGCGCTTCCAGGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 4408
Qy 289 AspMetGlnArgGlnTrpAlaGlyLeuValGluLysValGlnAlaAlaValGlyThrSer 308
Db 4409 GACATGCAAGCGCCAGTGGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 4468
Qy 309 AlaAlaProValProSerAspAsnHis 317
Db 4469 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4495
```

Search completed: March 14, 2003, 17:32:05
Job time : 2316.23 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581

Sequence: 1 MKVLMAALVFLAGCOAKV.....VEKVAAGVTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPPO.spool/US09827854/runat_11032003.101609.27476/app.query.fasta.1.3576
-DB=N.Geneseq.101002 -QFMT=fastlap -SUFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR.score=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854 -ECGN.1.1_1201 -runat_11032003.101609.27476 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_WMAP -LARGEQUERY -NEGSCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELext=7

Database :
1: N.Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	22 AAF84316	Human ApoE4 coding
2	1581	100.0	1156	24 AAD22047	Human apolipoprote
3	1573	99.5	954	24 AAD26035	Human apolipoprote
4	1573	99.5	1110	7 AAN60409	Human apolipoprote
5	1573	99.5	1147	22 ABA83113	Apolipoprotein E o
6	1573	99.5	1156	22 AAF84315	Human ApoE3 coding
7	1573	99.5	1156	24 AAD22048	Human apolipoprote
8	1569	99.2	1156	24 AAD22052	Human apolipoprote
9	1565	99.0	1156	22 AAF84314	Human ApoE2 coding
10	1565	99.0	1156	24 AAD22049	Human apolipoprote
11	1565	99.0	1156	24 AAD22051	Human apolipoprote
12	1564	98.9	1157	17 AAT06957	Human apolipoprote
13	1564	98.9	1157	24 ABN95746	Gene #2244 used to
14	1564	98.9	1157	24 ABK64514	Human benign prost
15	1564	98.9	1157	24 ABL65450	Lung cancer relate
16	1563	98.9	1110	6 AAN50450	Sequence encoding
17	1558	98.5	1116	24 AAD22050	Human apolipoprote
18	1469.5	92.9	1279	22 AAS22437	Human cDNA encodin
19	1450.5	91.7	1107	19 AAX75756	Human apolipoprote
20	1384.5	87.6	9360	24 ABL31915	Human ApoE genomic
21	1384.5	87.6	10716	24 AAD26034	Human apolipoprote
22	1379	87.2	3805	20 AAO09524	Human Apo E genom
23	1379	87.2	3805	20 AAO09526	Human Apo E genom
24	1370.5	86.7	10716	24 AAD26108	Human apolipoprote
25	1151.5	72.8	965	24 AAD32081	Bovine ApoE gene.
26	1151.5	72.8	5617	24 AAD32077	Human alpha-1-anti
27	1151.5	72.8	6026	24 AAD32075	Human albumin prom
28	1146	72.5	1126	19 AAV29159	Nucleotide sequenc
29	993	62.8	936	15 AAO69101	ApoE4Lx2 protease
30	993	62.8	936	17 AAT18070	Human ApoE4Lx2 CDN
31	972	61.5	660	18 AAT69792	Partial human apol
32	965	60.4	597	17 AAT18068	Human ApoE4 cDNA.
33	946	59.8	597	15 AAO69099	ApoE4L protease cd
34	872	55.2	1381	22 AAS22673	Human cDNA encodin
35	867.5	54.9	786	21 AAF18114	Lung cancer associ
36	686	43.4	600	20 AAV89595	EST clone CP147.
37	651	41.2	407	24 AABK34238	Human cDNA for nov
38	625	39.5	478	24 ABN95994	Colon adenocarcino
39	625	39.5	478	24 ABL67340	Thyroid cancer rel
40	594	37.6	499	24 AAS05049	Human diagnostic a
41	522	33.0	330	12 AAO011980	Human apolipoprote
42	504	31.9	405	21 AAO02139	Human secreted pro
43	485	30.7	345	22 AAH98479	Human EST-derived
44	485	30.7	345	22 AAH98479	Human EST-derived
45	410	25.9	382	21 AAA44191	Mouse secreted exp

ALIGNMENTS

RESULT 1
AAAF84316
AAAF84316 standard; cDNA: 1156 BP.

AAAF84316;
21-JUN-2001 (first entry)

Human ApoE4 coding sequence.

Human; ApoE4; Alzheimer's disease; arteriosclerosis; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 61..1014
/tag= a
/product= "Human ApoE4"

PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia -
 XX
 PS Claim 14; Page 80-81; 91pp; English.

CC The present sequence is a human apolipoprotein E (apoE)
 CC isoform, apolipoprotein E. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or preventing atherosclerosis without inducing
 CC hypertriglyceridemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.

XX Sequence 1156 BP; 208 A; 369 C; 432 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 4,64e-112 Length: 1156
 Score: 1581.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-827-854-14 (1-317) x AAD22047 (1-1156)

OY 1 MetlyValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGCGCGCTGCTGCTCATCTCTGCGAGATGCCAGGCCCAAGGTG 120
 OY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCTGCCAGACAGACAGCCAGTGGCAGAGCC 180
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCCAGCGCTGGAGACCTGGCACTGGTGGCTTTGGATTCACCTGCGGTGGCAGACA 240
 OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGCCAGCAACTGAGGGCG 300
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu 100
 DB 301 CTGATGAGCAGACCATATAGAGAGTTGAAGCCTTCAATGCGAATGAGAGACAACTG 360
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGACAGCGGGCGACGGCTGTCAGAGAGCTCAGCGCGCAGGCC 420
 OY 121 ArgLeuGlnAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGlnVal 140
 DB 421 CGCGTGGCGGAGACATGAGAGACCTGCGCGCCCTGCTGCTGCTCCACCTGCGG 480
 OY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGGCCAGACAGCCAGAGAGCTGGGGGCTGCTGCTCCACCTGCGC 540
 OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspSerGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGGCTAAGCGGCTCTCCGGCATGCTGATGACCTGCAAGAGCGCTGGCAGTATC 600
 OY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
 DB 601 CAGCGCGGGGGCGGAGGCGCGCAGCGCGCTCAGCGCGCATCCGAGCGCTGCGGG 660
 OY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGGTGGAGACAGGGCGGCTGGCGCCACTGTGGCTCCCTGCGCGCCAGCGCG 720
 OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGluMetLys 240

DB 721 CTACAGAGCGGGCCAGCGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
 OY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnValAlaGlnValArgAlaLys 260
 DB 781 AGCCGAGCCCGGACCGCTGAGCAGGTGAAGGAGCGAGTGGCGGAGTCCGCCCAAG 840
 OY 261 LeuGlnGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGAGAGAGAGAGCCCGACAGATACGCTGCAAGCGCGAGCGCTTCCAGCGCCCTCAAG 900
 OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGGTTCGAGCCCGCTGAGAGACATGCAGCGCCAGTGGCGCGGCTGTTGAGAGAAG 960
 OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTCCAGGCTGCGGTGGCAGCAGCGCCCGCTGTGCCAGCAGACAAATCAC 1011
 RESULT 3
 AAD26035
 ID AAD26035 standard; CDNA: 954 BP.
 AC
 XX AAD26035;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human apolipoprotein E (APOE) cDNA.
 XX
 KW Human; antilipemic; neuroprotective; nootropic; genetic variant; APOE;
 KW apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
 KW genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
 KW atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FT 1..954
 FT CDS
 FT
 FT /tag= a
 FT /product= "Human APOE protein"
 FT replace (13, C)
 FT
 FT variation
 FT
 FT /tag= a
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (31, G)
 FT
 FT variation
 FT
 FT /tag= b
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (364, A)
 FT
 FT variation
 FT
 FT /tag= c
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (388, C)
 FT
 FT variation
 FT
 FT /tag= d
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (498, G)
 FT
 FT variation
 FT
 FT /tag= e
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (526, T)
 FT
 FT variation
 FT
 FT /tag= f
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT replace (622, A)
 FT
 FT variation
 FT
 FT /tag= g
 FT /standard_name= "Single nucleotide polymorphism (SNP)"
 FT
 PN WO200179234-A2.
 XX
 XX 25-OCT-2001.
 PD
 XX
 XX 16-APR-2001; 2001WO-US12303.
 PF
 XX
 XX 14-APR-2000; 2000US-197188P.
 PR
 XX
 XX (GENA-) GENA/ISSANCE PHARM INC.
 PA
 XX
 XX Choi JY, Klem SE, Koshy B, Lee HH;

Alignment Scores:

Pred. No.: 1,81e-111
 Score: 1573.00
 Percent Similarity: 99.68%
 Best Local Similarity: 99.68%
 Query Match: 99.49%
 DB: 7

Length: 1110
 Matches: 316
 Conservative: 0
 Mismatches: 1
 Indels: 0
 Gaps: 0

US-09-827-854-14 (1-317) x AAN60409 (1-1110)

OY 1 MetlyValleuThrpAlaAlaLeuValThrpheLeuAglCysGlnAlaLysVal 20
 DB 15 ATGAAGTCTCTGTGGGCTGCGCTTGGTCAATCTCTGGCAGATGCGAGCCAAAGTG 74
 OY 21 GUGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
 DB 75 GAGCAGAGGGGTGGAGACAGAGCGGAGCCGAGCTGCCAGCACAGCCAGAGTGGCAGAGC 134
 OY 41 GylGlnATGTPGluLeuAlaLeuGlyArgPheTPAPSYrLeuAArgTPValGlnThr 60
 DB 135 GGCCAGCGCTGGAACTGGCACTGGGTGCTTTGGATTACCTGCTGGGTGGCAGACA 194
 OY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnLeuAArgAla 80
 DB 195 CTGTCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCAGTCCAGCAGAACTGAGGGCG 254
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 255 CTGATGAGCAGAACCATGAAAGAGTTGAAGGCTTCAAAATCGAAGCTGAGCAACACTG 314
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 315 ACCCGGTGGCGGAGAGACCGGGGCGAGCGCTGTCCAAAGAGACTGAGCGCGAGGCC 374
 OY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 375 CGGCTGGCGCGCCAGATGAGAGAGCTGTGCGCCGCTGTGTGCAATGCCGCGAGAGTG 434
 OY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuAArgValArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCGCATGCTGGCCAGACAGCAGGAGCTGCGGGGCGCTGCTCCACACTGGCGC 494
 OY 161 LysLeuAArgLysArgLeuLeuAArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 495 AAGCGCGCTAAGCGGCTCTCCGCCATGCGCATGACTGCGAGAGCGCTGGCACTGTAC 554
 OY 181 GlnAlaGlyAlaAArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 555 CAGGCGGGGCGCGGAGGGCGCCGAGCGGGCGCTCAGCGCCATCGCGAGGCGCTGGGG 614
 OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 DB 615 CCCCTGGGTGAACAGGGCGCGCTGGGGCGCGCACTGTGGGTCCCTGGCGGCGAGCGG 674
 OY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyLysArgLeuAArgAlaArgMetGluGlnMetGly 240
 DB 675 CTACAGAGAGCGGGCGCCAGGCGCTGGGGCGAGCGCGCTGCCGATGGAGAGATGGGGC 734
 OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 735 AGCCGAGACCCCGACCGCTGGAGAGGTGAAGGAGAGAGTGGGCGAGGTGGCGCCCAAG 794
 OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 795 CTGGGAGAGAGCGCCACAGATACGCTGCAAGCGCGAGGCGCTTCCAGGCGCGCTCAAG 854
 OY 281 SerTPPheGluProLeuValGluAspMetGlnArgGlnTPPALaglyLeuValGluLys 300
 DB 855 AGCTGTTGAGGCCCTGCTGTGAACATGCAGCCCACTGGGGCGGGCTGCTGTGAGAGAG 914
 OY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 915 GTGCAGCGTGGCGTGGGACCAAGCGCCCGCTGTGGCCAGCAGCAATCAC 965

RESULT 5
 ID ABA83113
 ID ABA83113 standard; DNA; 1147 BP.
 AC ABA83113;
 XX 08-FEB-2002 (first entry)
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
 XX
 DE Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN M0200175177-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 03-APR-2001; 2001WO-US10947.
 PF
 XX 03-APR-2000; 2000US-1943*CP.
 PR
 XX (US\$H) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PA Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 P1 WPI: 2001-626450/72.
 XX
 DR P-PSDB: ABB50287.
 XX
 PT Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene -
 XX
 PS Claim 23; Page 105-106; 140pp; English.
 XX
 CC The invention relates to methods for diagnosing and prognosing ovarian
 CC tumours in an individual via the detection and measurement of the
 CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer, in prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer and in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour. The ovarian tumour marker genes of the invention were identified
 CC using SAGE (serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of

CC the invention.

XX Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:

Pred. No.:	1,88e-111	Length:	1147
Score:	1573.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.49%	Indels:	0
DB:	22	Gaps:	0

US-09-827-854-14 (1-317) x ABA83113 (1-1147)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 46 ATGAAGTTCTGTGGCTGTGGCTGTGCATCTTCTGGCAGAGATCCAGGCCAAGGTG 105

QY 21 GluGlnAlaValAluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 106 GACCAACCGGTGGAGACAGACCGGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 165

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPValGlnThr 60
 DB 166 GGCACGCGCTGGAGACTGGCACTGGCTCTTTGGATTACCTGCGCTGGTGACAGACA 225

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 226 CTGCTCAGCAGGTGGAGAGAGAGCTGCTCAGCTCCAGGTCCAGCCAGGAAGTGGAGCGC 285

QY 81 LeuMetLysPcuThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 DB 286 CTGATGGACAGACCAAGAGAGTTGAAGCTTACAAATCGGAACCTGGAGAACACTG 345

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 346 ACCCGCGTGGCGAGAGACCGGCGCGCTGTCCAGAGAGTGGCGCGCGCAGGCC 405

QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 DB 406 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGGCGATCCCGCGGAGAGTG 465

QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 466 CAGGCCATGTCTGGCCAGACACCGAGAGACTCGGGGTGGCTCCGCTCCACCTGCGC 525

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnTyrArgLeuAlaValTyr 180
 DB 526 AAGCTGGTAAGGCGCTCTCCCGATGCCGATGACCTGCAGAAAGCCCTGGCAGATGAC 585

QY 181 GlnAlaGlyAlaArgGluGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 586 CAGGCGGCGCGCGCGAGCGCGCGAGCGCGCTCCAGCCATCCGCGAGCCCTGGGG 645

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 646 CCCCTGGTGAACAGGCGCGCTGGCGCGCGCACTGTGGCTCCCTGGCGCGCGCAGCCG 705

QY 221 LeuGlnGluArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 DB 706 CTACAGAGAGGCGGCCAGGCTGGGCGAGCGGCTGCGCGCGCGGATGGAGAGATGGGC 765

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
 DB 766 AGCGGAGACCGCGCGCGCTGGAGAGGTGAAGAGAGAGTGGCGGAGGTGGCGCGCAAG 825

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 826 CTGGAGAGACGAGCGCCAGAGATACCTGCGAGCGAGCGCTTCCAGGCGCCGCTCAAG 885

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 886 AGCTGTTCGAGCCCTCTGTGGAGACATGCACGCGCAGTGGCGCGGCTGTGTGGAGAG 945

QY 301 ValGlnAlaAlaValAluThrSerAlaAlaProValProSerAspAsnHis 317
 DB 946 GTGCAGGCTGCGGTGGGACAGGCGCGCTGTGGCCAGCAGCATCTAC 996

RESULT 6

AAAF84315

ID AAF84315 standard; cDNA; 1156 BP.

XX AAF84315;

XX 21-JUN-2001 (first entry)

XX Human ApoE3 coding sequence.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 61..1014

XX FT /tag= a

XX JP2001017028-A.

XX 23-JAN-2001.

XX 28-APR-2000; 2000JP-0128919.

XX 06-MAY-1999; 99JP-0125647.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI: 2001-285406/30.

XX P-PSDB; AAB80997.

XX New apoE humanized mammalian cell useful for screening for agents

XX useful for treating or preventing Alzheimer's disease and

XX arteriosclerosis -

XX PS Disclosure; Page 13-14; 22pp; Japanese.

XX CC The present invention relates to an ApoE humanised mammalian cell. The

XX CC present sequence is the coding sequence for human ApoE3, which was used

XX CC in the method of the present invention. The ApoE humanised mammalian cell

XX CC can be used for screening for agents useful for treating or preventing

XX CC Alzheimer's disease and arteriosclerosis.

XX SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

US-09-827-854-14 (1-317) x AAF84315 (1-1156)

Alignment Scores:

Pred. No.:	1.9e-111	Length:	1156
Score:	1573.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.49%	Indels:	0
DB:	22	Gaps:	0

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGTTCTGTGGCTGTGGCTGTGCATCTTCTGGCAGAGATCCAGGCCAAGGTG 120

QY 21 GluGlnAlaValAluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GACCAACCGGTGGAGACAGACCGGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPValGlnThr 60
 DB 181 GGCACGCGCTGGAGACTGGCACTGGCTCTTTGGATTACCTGCGCTGGTGACAGACA 240

QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerGIuValThrGIuLeuArgAla 80
 DB 241 CTGTCTGACAGAGTCAGAGAGAGCTGCTCCAGCTCCAGAGTCAACCCAGAACTGAGGCG 300
 QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
 DB 301 CTGATGGAGACATGATGAAGAGTTGAAGGCTTCAAAATCGAACTGAGAGAACACTG 360
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTCAGCGCGCAAGCC 420
 QY 121 ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTyrArgGIuVal 140
 DB 421 CGGCTGGCGCGGACATGAGAGACTGTGGCCCGCTGTGTGCTACCTACCGCGGCAAGTG 480
 QY 141 GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTCGGCACAGACACCGAGAGCTGCGGTGCTGCTCCACCTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
 DB 541 AAGCTGCTGACGCTCTCTCCGATGCGCATGACCTGCAAGAGCGCTGGAGTGTAC 600
 QY 181 GIuAlaGIuAlaArgGIuGIuValArgGIuLysSerAlaAlaArgGIuArgLeuGIu 200
 DB 601 CAGCGCGGGCGCGGAGGCGCGGCGCGCTCAGCGCATCGCGGAGCGCTGGGG 660
 QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuPro 220
 DB 661 CCCCTGTGTGAAACAGGCGCGCTGCGGCGCGCTGTGTGCTCTCTGCGGCGAGCG 720
 QY 221 LeuGIuGIuArgAlaGIuAlaArgGIuGIuArgLeuArgAlaArgMetGIuGIuMetGIu 240
 DB 721 CTACAGAGAGCGGCGGAGGCTGTGGGCGAGCGGCTGCGCGGATGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuValArgAlaLys 260
 DB 781 AGCGGAGCGCGGAGCGCTGTGAGAGAGTGAAGAGAGAGTGGCGGAGCTGCGGCCAAG 840
 QY 261 LeuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIuGIu 280
 DB 841 CTGGAGGAGAGCGGCGGAGGCTGAGTACGCTGCAGGCGGAGGCTTCCAGCGCGCTCAAG 900
 QY 281 SerTrpPheGIuProLeuValGIuAspMetGIuArgGIuIntTrpAlaGIuValGIuLys 300
 DB 901 AGCTGTGTGAGCGGCTGTGTGAGAGCATTCAGCGCCAGTGGGCGGCTGTGTGAGAG 960
 QY 301 ValGIuAlaAlaValGIuThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTGCGGTGGGACACGCGCGCCCTGTGTGCCAGGACAAATCAC 1011
 RESULT 7
 AAD22048 ID AAD22048 standard; DNA; 1156 BP.
 AC AAD22048;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 XX
 KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 61..1014
 FT /*tag= a
 FT /product= "Human apoE isoprotein, apoE3"
 FT sig_peptide 61..114
 FT /*tag= b

FT mat_peptide 115..1011
 FT /*tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 XX
 XX WO200177136-A1.
 XX
 XX 18-OCT-2001.
 XX
 XX 06-APR-2001; 2001WO-US11358.
 XX
 XX 06-APR-2000; 2000US-0544386.
 XX
 XX 04-OCT-2000; 2000US-0679088.
 XX
 XX 05-APR-2001; 2001US-0827854.
 XX
 XX (KOSP-) KOS PHARM INC.
 XX (UYBO-) UNIV BOSTON.
 XX
 XX Zannis VI, Kyreos KE;
 XX
 XX WPI: 2002-010885/01.
 XX
 XX P-PSDB; AAE13294.
 XX
 XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia
 XX
 XX Claim 14; Page 81; 91pp; English.
 XX
 XX The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
 SO
 Alignment Scores:
 Pred. No.: 1.9e-111 Length: 1156
 Score: 1573.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best local Similarity: 99.68% Mismatches: 1
 Query Match: 99.49% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-14 (1-317) x AAD22048 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGIuGIuGIuAlaLysVal 20
 DB 61 ATGAAGTTCTGTGGGCTGCGTGTGCTGCATTCCTGCGAGAGATGCCAGCAAGGTG 120
 QY 21 GIuGIuAlaValAlaGIuThrGIuProGIuProGIuLeuArgGIuGIuIntTrpGIuIns 40
 DB 121 GAGCAAGCGGTGGAGACAGACCGGAGCGGAGCGCGCCAGCGAGCAAGCGATGGAGAGC 180
 QY 41 GIuGIuArgTrpGIuLeuAlaLeuGIuLysArgPheTrpAspTyrLeuArgTrpValGIuIns 60
 DB 181 GGCAGGCGCTGGAACTGGCACTGGGTGCTTTGGGATACCTGCGCTGGGTGAGACA 240
 QY 61 LeuSerGIuGIuValGIuGIuGIuLeuLeuSerGIuValIntTrpGIuLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGTCCAGGTCCAGGAGACTGAGGCG 300
 QY 81 LeuMetAspGIuTrpMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
 DB 301 CTGATGGAGACATGATGAAGAGTTGAAGGCTTCAAAATCGAACTGAGAGAACACTG 360
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGAGTGGCGGCGAGGCC 420

QY 121 ArgLeuGluGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGGCTGGGCGCGACATGAGGAGACGTGCGCGCTGTGTGACGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGCCATGCTGCGCCAGACACCGAGAGACTCGGGTGGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 541 AACCTGCGTAAGGGGCTCTCCCGGATGCCGATGACCTGAGAAAGCCCTGGGAGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGly 200
|||||
Db 601 CAGCGCGGGCGCGGAGGGCGCGAGCGGGGCTCAGCGCCATCCGCGAGCGCGCTGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 661 CCCCTGTGTGAAACAGGGCCCGCTGCGGGCCGCACTGTGGCTCCCTGGCGCGCACGCC 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrPglGlyLeuArgLeuArgAlaArgMetGluGlnMetGly 240
|||||
Db 721 CTACAGAGAGCGGGCCCGGCTGGGGGAGCGGCTGCGCGCGGATGGAGAGATGGG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
|||||
Db 781 AGCGGAGCGCGGACCGCTGAGAGAGGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||
Db 841 CTGAGAGAGAGAGCGCCAGCATACGCTGACGCGGAGCGGCTCCAGAGCGCGCTCAAG 900
QY 281 SerTrpPheGlnProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 901 AGCTGCTGCAAGCCCTGGTGGAGAGACATCAGCGCCAGTGGGCGGGGCTGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 961 GTGACAGGCTGCGTGGGACACGCGCGCCCTGTGCGCCAGCACATATCAC 1011
RESULT 8
AAD22052
ID AAD22052 standard; DNA; 1156 BP.
XX
AC AAD22052;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX
KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT sig_peptide /tag= a
FT mat_peptide /product= "Human apoE allele, apoE2**"
FT 115..1011 /tag= b
FT /tag= c
FT /product= "Mature human apoE allele, apoE2**"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001: 2001WO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
DR WPI: 2002-010885/01.
DR P-PSDB: AAE13298.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridemia
XX
PS Claim 14; Page 83; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2** DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other:
Alignment Scores:
Pred. No.: 3,83e-111 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-14 (1-317) x AAD22052 (1-1156)
QY 1 MetLysValLeuTrpAlaIleLeuValIleThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 61 ATGAAAGGTTCTGTGGGCTGCGCTGTGCTGACATTCTCGGACAGATGCCAGGCCAAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
|||||
Db 121 GAGCAAGCGGTGGAGAGACAGAGCGCGAGCCGACCTCGCGCAGACAGCGAGTGCAGAGC 180
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 181 GGCACAGCGCTGGGAACTGGACAGGCGTCTTTGGGATTACTCGCGCTGGGTCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValIleThrGlnGlnLeuArgAla 80
|||||
Db 241 CTGTCTGACAGGTGACGAGAGAGCTGCTCAGCTCCAGGTACCCAGAACTGAGAGGCG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
|||||
Db 301 CTGATGAGAGACCATGAAGAGAGTGAAGCGCTCAAAATCGAACTGAGAGAACACTG 360
QY 101 ThrProValAlaGlnGlnIleThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
|||||
Db 361 ACCCGGCTGGCGAGAGAGAGCGCGGCTGTCCAAAGAGCTGACGCGCGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGGCTGGGCGCGACATGAGAGAGCTGTGCGCGCTGTGTGACGTACCGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGCCATGCTGCGCCAGAGACCGGAGAGCTGGGGTGGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||

CC	Alzheimer's disease and arteriosclerosis.
CC	The ApoE humanized mammalian cell useful for screening for agents
CC	in the method of the present invention. The ApoE humanized mammalian cell
CC	can be used for screening for agents useful for treating or preventing
CC	Alzheimer's disease and arteriosclerosis.
XX	
XX	Disclosure; Page 11-12; 22pb; Japanese.
XX	
XX	New apoE humanized mammalian cell useful for screening for agents
PT	useful for treating or preventing Alzheimer's disease and
PT	arteriosclerosis -
XX	
DR	P-PSDB; AAB80996.
XX	
XX	WPI; 2001-285406/30.
XX	
PA	(MITU) MITSUBISHI CHEM CORP.
XX	
PR	06-MAY-1999; 99JP-0125647.
XX	
XX	
PN	JP2001017028-A.
XX	
PD	23-JAN-2001.
XX	
PF	28-APR-2000; 2000JP-0128919.
XX	
FT	CDS
FT	Location/Qualifiers 61..1014 /*tag= a /product= "Human ApoE2"
XX	
OS	Homo sapiens.
XX	
KW	Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX	
DE	Human ApoE2 coding sequence.
XX	
DT	21-JUN-2001 (first entry)
XX	
AC	AAEF84314;
XX	
ID	AAF84314 standard; cDNA; 1156 bp.
RESULT 9	

QY	1	MethylsValLeuThrPalaIalaIeuIeuValThrPheIeuAlaGlyCysGlnAlaIaIysVal	20
Db	61	ATGAAGGTTCTGTGGGCTGGCTGCTGTGTATCAATTCCTGGCAAGATGTCCAGGCCAAGGTG	120
QY	21	GIuGlnAlaIaValGIuThrGIuProGIuLeuPheIeuAArgGlnGlnThrGIuTProIIns	40
Db	121	GAGCAAGGGGTGGAGACAGAGAGCCGAGGCCGAGCTGGCCAGCAGACCAGAGTGGAGAGC	180
QY	41	GlyGlnAArgTProGIuLeuAlaLeuGlyAArgPheTProPaspTyrLeuAArgTProValGlnThr	60
Db	181	GGCCAGCCCTGGGAACCTGGCACTGGGTGGCTTTGGGTATCACTTCGCTGGGTGCAGACA	240
QY	61	LeuSerGIuGlnIValGIuGlnGluLeuIeuIeuSerSerGIuAlaThrGlnGluLeuAArgIa	80
Db	241	CTGTCTGACAGAGTCCAGAGAGAGCTGCTCACTCCAGAGTACCCAGAGAACCTGAGGGCG	300
QY	81	LeuMetAspGIuThrMetIysGIuLeuIysAlaTyrIlySerGIuLeuGluGluGlnLeu	100
Db	301	CTGATGGACAGAACCATTAAGAGATTGAAGGCTCTCAATCGAATCGAAGCTGAGAACCACTG	360
QY	101	ThrProValAlaGluGluThrAArgAlaAArgIeuSerIlySGluLeuGlnAlaIaGlnAla	120
Db	361	ACCCGGTGGCGGAGAAACGGGGCCAGCGCTGTCCAAGAGCTGACAGCGGAGGCCAGGCC	420
QY	121	AArgLeuGIyAlaAspMetGIuAspAlaIArgGIyAArgIeuAlaGlnTyrAArgIyGIuVal	140
Db	421	CGGCTGGGGCGCGACATGAGAGACTGTGGCGCGCTCGTGGTCACTACCGGGCGGAGGTG	480
QY	141	GlnAlaMetLeuGlyGlnSerThrGIuGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160
Db	481	CAGGCCATGCTCGCGCAAGACACCAGAGAGCTGGCGGTGCGCTCCCTCCACCTGGCGC	540
QY	161	LysIeuAArgIysAArgLeuIeuAArgAspAlaAspAspIeuGlnIlyAArgLeuAlaValTyr	180
Db	541	AAGCTGCGTAAAGGGCTCTCCGCGATCCGATGACTTCGACAGAACTGGCTTGGCACTGTAC	600
QY	181	GlnAlaGIyAlaAArgGIuGlyAlaGluAArgIyIeuSerAlaIleAArgGIuAArgLeuGIy	200
Db	601	CAGGCCGGGGCCCCGGAGGGGGCCGAGCGGGCCACAGCCGCAATCCGGAGACGGCTGGGG	660
QY	201	ProIeuValGIuGlnGlyAArgValAArgAlaIaThrValIlySerIleuAlaGlyGlnPro	220
Db	661	CCCCGTGTGACAAGAGGGCGGGTGGCGGCCCGCCACTGTGGGCTCCCTGGCGGGCAGCGG	720
QY	221	LeuGlnGIuAArgAlaGlnAlaIaTProIyGIuAArgLeuAArgAlaAArgMetGIuGluMetGIy	240
Db	721	CTACAGGAGAGGGGCCAGAGGCTGGGGGGCAGAGGGCTGGCGCGCGGATGAGAGATGGCC	780
QY	241	SerAArgThrAArgAspAArgIeuAspGIuValIlySGluGlnIValAlaGluValAArgAlaIys	260
Db	781	AGCCCGACCCCGAGACCGCTGTGACAGAGGTGAAGAGCGAGGTGGGAGGTCCGGCCAGAG	840
QY	261	LeuGIuGIuGlnAlaGlnGlnIleArgIeuGlnAlaGlnAlaPheGlnAlaAArgLeuIys	280
Db	841	CTGGAGAGAGAGCCCGACAGATACGCTTCGAGAGCGCAGAGGCTTCCAGGCCCGGCTCAAG	900
QY	281	SerTProPheGIuProLeuValGIuAspMetGIuAArgIuTProAlaGlyLeuValGIuIlys	300
Db	901	AGCTGGTTCGAGCCCTGTGTGAAGACTGTGCAGGCCAGTGGCGGGCTGTGTGGAGAG	960

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACACGCGCCCTGTGCGCCAGCAATCACC 1011
RESULT 10
AAD22049
ID AAD22049 standard; DNA; 1156 BP.
XX
AC AAD22049;
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT sig_peptide 61..114 /product= "Human apoE isoprotein, apoE2"
FT mat_peptide 115..1011 /tag= b
FT /tag= c /product= "Mature human apoE isoprotein, apoE2"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNITV BOSTON.
XX
PI Zannis VI, Kypreos KE;
XX
DR WPI: 2002-010885/01.
DR P-PSDB: AAE13295.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia
XX
PS Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 7.75e-111 Length: 1156
Score: 1565.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.99% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-14 (1-317) x AAD22049 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGGCTGTGGTGTGTCACATTCTCGCAGAGATGCCAGGCCAAGGTT 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGCCAGCAGCAGCCGAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCACGCGCTGGGAGACTGGGCTGGCTTGGGATTTACCTGGCTGGTGGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTACGCTCCAGGTCCAGGAACTGAGAGGGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
Db 301 CTGATGGAGCAGACCATGAAGAGTTGAAGGCCCTACAAATCCGAACTGGAGGAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGCTGGCGGAGAGACGCGGCGACGCTCTCCAGAGCTGAGGCGGCGCAGGCC 420
QY 121 ArgLeuGluValAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGGAGAGCGTGTGGCGGCTGTGGTGCAGTACCGCGGCGAGGTT 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGCAGAGACCGAGAGCTGCGGCTCGCTCCCTCCACCTGCCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaLysTrp 180
Db 541 AACCTCGTGAAGCGGCTCTCCGCGATGCCGAGTCTCAGAAAGCTCGGAGAGTGTAC 600
QY 181 GlnAlaGluValAlaArgGluGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCCGGGGCCCGGACAGGCGCGGAGCGGCGCTCAGCGCCATCCGACAGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGGAACAGAGGCGCGGCGGCGGCGCAGCTGTGGGCTCCCTGGCCGCGCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGCGGGCCAGGCGCTGGGCGAGCGGCTGCGGCGGAGTGAAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnGluValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCCCGGCGGCGGCGCTGAGCGAGTGAAGAGCAGTGGCGGAGGCGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAAGACAGGCGCCAGAGATATCCCTGCAAGCGGCGGAGCCCTTCAGAGCCCGCCCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGCTTGAAGCCCTGTGGTGAAGCATGACAGCCCACTGGGCGGCTGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACACGCGCCCTGTGCGCCAGCAATCACC 1011
RESULT 11
AAD22051
ID AAD22051 standard; DNA; 1156 BP.
XX
AC AAD22051;
XX
DT 12-FEB-2002 (first entry)

XX	Human apolipoprotein E (apoE) allele, apoE2* DNA.
DE	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
XX	hypertriglyceridaemia; low density lipoprotein; LDL; ds.
KW	Homo sapiens.
RW	
XX	
OS	
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	61..1014
FT	/tag= a
FT	/product= "Human apoE allele, apoE2*"
FT	sig_peptide
FT	61..114
FT	/tag= b
FT	mat_peptide
FT	115..1011
FT	/tag= c
FT	/product= "Mature human apoE allele, apoE2*"
PX	
PN	WO200177136-A1.
XX	
PD	18-OCT-2001.
XX	
PE	06-APR-2001; 2001MO-US11358.
XX	
PR	06-APR-2000; 2000US-0544386.
PR	04-OCT-2000; 2000US-0679088.
PR	05-APR-2001; 2001US-0827854.
XX	
PA	(KOSP-) KOS PHARM INC.
XX	(UYBO-) UNIV BOSTON.
PI	Zannis VI, Kypros KE.
DR	WPI: 2002-010885/01.
XX	P-PADB: AAE13297.
PT	New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT	cholesterol, delaying the onset of or treating atherosclerosis in
PT	mammal, without inducing hypertriglyceridemia
PS	Claim 14; Page 82; 91pp; English.
XX	
CC	The present sequence is a human apolipoprotein E (apoE)
CC	allele, apoE2* DNA. The apoE lipoproteins are useful for
CC	lowering cholesterol, delaying the onset of atherosclerosis,
CC	treating or regressing atherosclerosis without inducing
CC	hypertriglyceridaemia, in a mammal lacking an endogenous,
CC	normally functioning apoE gene or low density lipoprotein (LDL)
CC	receptor or is at risk for developing atherosclerosis due to
CC	accumulation of lipoprotein remnants in the bloodstream or having
CC	a defect in remnant removal.
XX	
SQ	Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:	
Pred. No.:	7.75e-111
Score:	1565.00
Percent Similarity:	99.37%
Best local Similarity:	99.37%
Query Match:	98.99%
DB:	24
US-09-827-854-14 (1-317) x AAU22051 (1-1156)	
OY	1 MettysValLeuTPPAlaIaleuLeuValThrPhenuAlaglyGscglnAlayVal 20
Db	61 ATGAAGGTCTGTGGCTGGCTGTCGTACATCTCGCAGAAGTCCAGGCCAAGGTG 120
OY	21 GluGlAlaValGIuThrGIuProGLILProGLIUdeuArgInglInThrGIuTPGINser 40
Dd	121 GAGCAAGCGGTGGAGACAGAGCCGCAGCCGAGCTGGCCAGCAGACCGAGTGCAAGC 180
OY	41 GlyAlnArGrTrpGIuLeuAlaIaleuGlyArGrPheTrpAspTyrlenuArGrTrpValGIuThr 60

	Accession	Gene	Species	Length (bp)	Start (bp)	End (bp)	Strand	GC Content (%)	GC Skew	GC Bias	GC Bias2	GC Bias3	GC Bias4	GC Bias5	GC Bias6	GC Bias7	GC Bias8	GC Bias9	GC Bias10	GC Bias11	GC Bias12	GC Bias13	GC Bias14	GC Bias15	GC Bias16	GC Bias17	GC Bias18	GC Bias19	GC Bias20	GC Bias21	GC Bias22	GC Bias23	GC Bias24	GC Bias25	GC Bias26	GC Bias27	GC Bias28	GC Bias29	GC Bias30	GC Bias31	GC Bias32	GC Bias33	GC Bias34	GC Bias35	GC Bias36	GC Bias37	GC Bias38	GC Bias39	GC Bias40	GC Bias41	GC Bias42	GC Bias43	GC Bias44	GC Bias45	GC Bias46	GC Bias47	GC Bias48	GC Bias49	GC Bias50	GC Bias51	GC Bias52	GC Bias53	GC Bias54	GC Bias55	GC Bias56	GC Bias57	GC Bias58	GC Bias59	GC Bias60	GC Bias61	GC Bias62	GC Bias63	GC Bias64	GC Bias65	GC Bias66	GC Bias67	GC Bias68	GC Bias69	GC Bias70	GC Bias71	GC Bias72	GC Bias73	GC Bias74	GC Bias75	GC Bias76	GC Bias77	GC Bias78	GC Bias79	GC Bias80	GC Bias81	GC Bias82	GC Bias83	GC Bias84	GC Bias85	GC Bias86	GC Bias87	GC Bias88	GC Bias89	GC Bias90	GC Bias91	GC Bias92	GC Bias93	GC Bias94	GC Bias95	GC Bias96	GC Bias97	GC Bias98	GC Bias99	GC Bias100	GC Bias101	GC Bias102	GC Bias103	GC Bias104	GC Bias105	GC Bias106	GC Bias107	GC Bias108	GC Bias109	GC Bias110	GC Bias111	GC Bias112	GC Bias113	GC Bias114	GC Bias115	GC Bias116	GC Bias117	GC Bias118	GC Bias119	GC Bias120	GC Bias121	GC Bias122	GC Bias123	GC Bias124	GC Bias125	GC Bias126	GC Bias127	GC Bias128	GC Bias129	GC Bias130	GC Bias131	GC Bias132	GC Bias133	GC Bias134	GC Bias135	GC Bias136	GC Bias137	GC Bias138	GC Bias139	GC Bias140	GC Bias141	GC Bias142	GC Bias143	GC Bias144	GC Bias145	GC Bias146	GC Bias147	GC Bias148	GC Bias149	GC Bias150	GC Bias151	GC Bias152	GC Bias153	GC Bias154	GC Bias155	GC Bias156	GC Bias157	GC Bias158	GC Bias159	GC Bias160	GC Bias161	GC Bias162	GC Bias163	GC Bias164	GC Bias165	GC Bias166	GC Bias167	GC Bias168	GC Bias169	GC Bias170	GC Bias171	GC Bias172	GC Bias173	GC Bias174	GC Bias175	GC Bias176	GC Bias177	GC Bias178	GC Bias179	GC Bias180	GC Bias181	GC Bias182	GC Bias183	GC Bias184	GC Bias185	GC Bias186	GC Bias187	GC Bias188	GC Bias189	GC Bias190	GC Bias191	GC Bias192	GC Bias193	GC Bias194	GC Bias195	GC Bias196	GC Bias197	GC Bias198	GC Bias199	GC Bias200	GC Bias201	GC Bias202	GC Bias203	GC Bias204	GC Bias205	GC Bias206	GC Bias207	GC Bias208	GC Bias209	GC Bias210	GC Bias211	GC Bias212	GC Bias213	GC Bias214	GC Bias215	GC Bias216	GC Bias217	GC Bias218	GC Bias219	GC Bias220	GC Bias221	GC Bias222	GC Bias223	GC Bias224	GC Bias225	GC Bias226	GC Bias227	GC Bias228	GC Bias229	GC Bias230	GC Bias231	GC Bias232	GC Bias233	GC Bias234	GC Bias235	GC Bias236	GC Bias237	GC Bias238	GC Bias239	GC Bias240	GC Bias241	GC Bias242	GC Bias243	GC Bias244	GC Bias245	GC Bias246	GC Bias247	GC Bias248	GC Bias249	GC Bias250	GC Bias251	GC Bias252	GC Bias253	GC Bias254	GC Bias255	GC Bias256	GC Bias257	GC Bias258	GC Bias259	GC Bias260	GC Bias261	GC Bias262	GC Bias263	GC Bias264	GC Bias265	GC Bias266	GC Bias267	GC Bias268	GC Bias269	GC Bias270	GC Bias271	GC Bias272	GC Bias273	GC Bias274	GC Bias275	GC Bias276	GC Bias277	GC Bias278	GC Bias279	GC Bias280	GC Bias281	GC Bias282	GC Bias283	GC Bias284	GC Bias285	GC Bias286	GC Bias287	GC Bias288	GC Bias289	GC Bias290	GC Bias291	GC Bias292	GC Bias293	GC Bias294	GC Bias295	GC Bias296	GC Bias297	GC Bias298	GC Bias299	GC Bias300	GC Bias301	GC Bias302	GC Bias303	GC Bias304	GC Bias305	GC Bias306	GC Bias307	GC Bias308	GC Bias309	GC Bias310	GC Bias311	GC Bias312	GC Bias313	GC Bias314	GC Bias315	GC Bias316	GC Bias317	GC Bias318	GC Bias319	GC Bias320	GC Bias321	GC Bias322	GC Bias323	GC Bias324	GC Bias325	GC Bias326	GC Bias327	GC Bias328	GC Bias329	GC Bias330	GC Bias331	GC Bias332	GC Bias333	GC Bias334	GC Bias335	GC Bias336	GC Bias337	GC Bias338	GC Bias339	GC Bias340	GC Bias341	GC Bias342	GC Bias343	GC Bias344	GC Bias345	GC Bias346	GC Bias347	GC Bias348	GC Bias349	GC Bias350	GC Bias351	GC Bias352	GC Bias353	GC Bias354	GC Bias355	GC Bias356	GC Bias357	GC Bias358	GC Bias359	GC Bias360	GC Bias361	GC Bias362	GC Bias363	GC Bias364	GC Bias365	GC Bias366	GC Bias367	GC Bias368	GC Bias369	GC Bias370	GC Bias371	GC Bias372	GC Bias373	GC Bias374	GC Bias375	GC Bias376	GC Bias377	GC Bias378	GC Bias379	GC Bias380	GC Bias381	GC Bias382	GC Bias383	GC Bias384	GC Bias385	GC Bias386	GC Bias387	GC Bias388	GC Bias389	GC Bias390	GC Bias391	GC Bias392	GC Bias393	GC Bias394	GC Bias395	GC Bias396	GC Bias397	GC Bias398	GC Bias399	GC Bias400	GC Bias401	GC Bias402	GC Bias403	GC Bias404	GC Bias405	GC Bias406	GC Bias407	GC Bias408	GC Bias409	GC Bias410	GC Bias411	GC Bias412	GC
--	-----------	------	---------	-------------	------------	----------	--------	----------------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	----

```

FT      /tag= a
FT      note= "misc-signal"
FT      CDS 62..1015
FT      /tag= b
XX      US5472858-A.
XX      PD 05-DEC-1995.
XX      PF 04-JUN-1991: 91US-0709949.
XX      PR 04-JUN-1991: 91US-0709949.
XX      PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX      PI Attie AD, Beckage NE, Gretch DG, Sturley SL;
XX      DR WPI; 1996-029812/03.
XX      DR P-PSDB; AAR86791.
XX      PT Prod. of recombinant apolipoprotein E in insects - by infecting
XX      PT Manduca sexta larvae with recombinant Autographa californica
XX      PT nuclear polyhedrosis baculovirus vector.
XX      PS Disclosure; Columns 11-14; 10pp; English.
XX      CC Recombinant human apolipoprotein-E (APOE) (AAR86791) can be produced
XX      CC by preparing a genetic construct (contg. an APOE-encoding sequence,
XX      CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
XX      CC to be expressed in insect cells), which is then introduced into a
XX      CC Manduca sexta larva (using a recombinant Autographa californica
XX      CC haemolymph of the larval host. The APOE produced is in a form
XX      CC sufficiently complexed with lipids to be biologically active, which
XX      CC cannot be achieved in insect cell cultures, and can therefore be
XX      CC used in therapeutic applications.
XX      SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 9, 24e-111 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
DB: 17 Gaps: 0

US-09-827-854-14 (1-317) x AAT06957 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaTysVal 20
DB 62 ATGAAGATTCTCTGGGCTGGTGTGTCACATTCTCTGGCAGAGATGCCAGGCAAGT 121
QY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluP 40
DB 122 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGGCAGACAGCCAGTGGCAG 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValG 60
DB 182 GGCACGCGCTGGGAATCGCACCTGGCTTTGGGATTACCTGGCGTGGGTGGCAGA 241
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeu 80
DB 242 CTGTCTGACGACAGGTGACGAGAGCTGTCTGACGCTCCCAAGTCAACCAAGAC 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlu 100
DB 302 CTGATGGACGAGACCAATGTAAGAGTTGAAGGCTTACAAATCGGAATCGAGGA 361
QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGln 120
DB 362 ACCCCGCTGACGAGACGCGGCGGCTGTCTCCAAAGGAGCTGACAGACGGCGCAG 421

```

```

QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 422 CGGCTGGCGCGGACATGAGAGACGTGTGGGGCCGCTGCTGTGACGATACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTCGGCCAGAGACACCGAGAGCTCCGGGTGGCTCCCTCCCACTGCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AACCTCGTAAAGGCTCTCTCCGCAATCCGATGACCTGCAGAGCGCTTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 602 CAGCGCGGGGCCCGCGAGGGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCGCTCGGGG 661
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 662 CCCCTGTGTGAAACAGGCGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 722 CTACAGAGAGGGGCCCAAGGCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGCC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnAlaValArgAlaLys 260
DB 782 AGTCGAGACCGCGACCGCTGCGAGAGTGAAGAGACAGATGGCGGAGTGGCGCGCAAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 842 CTGAGAGACGAGCGCCAGAGATACCTGTCAGAGCGCGGCTTCCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB 902 AGCTGGTTGAGACCCCTGCTGTGAAGACATGACAGCGCACTGGCGCGCTGTGAGAAAG 961
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAGGCTGCGTGGGCGACAGCGCGCGCTGTGCTCCAGCGACATATCAG 1012

RESULT 13
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
AC ABN95746;
XX 13-AUG-2002 (first entry)
DT
DE Gene #2244 used to diagnose liver cancer.
XX
DE Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumour; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a

```

PT liver tissue sample -
XX
PS Claim 1: SEQ ID NO 2244; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 9 24e-111 Length: 1157
Score: 1564.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.92% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-14 (1-317) x ABN95746 (1-1157)
QY 1 MettysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaAlaVal 20
Db 62 ATGAAGGTTCTGTGGCGCTGCTGCTCACTTCCTGGCAGATGCGCAGCCAGGTG 121
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAGACGGTGGAGACAGACCCGAGCCGAGCTGCCAGACAGACCGATGGCAGAC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 182 GGCACGCGCTGGAACTGGCACTGGTGGCTTTGGGATTCACCTGGCTGGGTGAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGCTGTAGCAGGTGCAGGAGGAGCTGCTCACTCCCAAGTCCACCCAGAACTGAGGCG 301
QY 81 LeuMetAspGluThrMetLeuGluLeuLysAlaTrpTrpLysSerGluLeuGluGlnLeu 100
Db 302 CTGATGAGCAGACCATGATGAGAGTGAAGGCTTACAAATCGCAACTGGAGAACAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGGAGACCGCGGCGCTGTCACAAAGAGCTGCACACGCGCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValArgGlnArgLeuValGlnTrpArgGlnVal 140
Db 422 CGGCTGGGCGCGGACATGGAGGAGCTGTGGCGGCCCTGTGTCAGTACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTCGGCGCAAGACACCGAGAGCTGGCGGGTGCCTTCGCTCCACCTGGCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
Db 542 AAGCTGCTAAGCGGCTCTCCGCGATCCGATACCTGCAGAAAGCGGCTGGCAGTGTAC 601
QY 181 GlnAlaGlnAlaArgGluGluGlnValAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 602 CAGGCGGGGCGCGGAGGCGCGCGGCGCTCAGCGCCATCCGAGAGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220

Db 662 CCCCTGGTGAACAGGCGCGGCGCGCCACTGTGGGCTCTGGCGCGCAGCCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 722 CTACAGGAGCGGCGCCAGGCTGGGGGCGAGGCTGCGCGCGGATGTGAGATGGCG 781
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTGGAGACCGCGACCGCTGTGAGAGTGAAGAGCAGGAGTGGGAGAGTCCGCGCCAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGGAGCAGCGCCACAGATACGCTGCAGCGCGGCGCTTCAGAGCCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGGTTCAGCCCTGCTGGAAAGACTGACAGCCAGTGGCGCGGCTGTGGAGAG 961
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGTGGCGACAGCGCGCCCTGTGGCCAGCAGCAATCAC 1012
RESULT 14
ABK64514
ID ABK64514 standard; DNA; 1157 BP.
AC ABK64514:
XX 18-JUN-2002 (first entry)
DT Human benign prostatic hyperplasia gene #409.
DE Human: benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.
XX WO200212440-A2.
XX 14-FEB-2002.
PD 07-AUG-2001; 2001WO-US24708.
PE 07-AUG-2000; 2000US-223323P.
XX 07-AUG-2000; 2000US-223323P.
PR 05-JUN-2001; 2001US-0873319.
XX (GENE-) GENE LOGIC INC.
PA (NISR) JAPAN TOBACCO INC.
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX detecting expression levels of one or more genes in prostate cells from
XX patient that are differentially regulated compared to normal prostate
XX cells -
PT
PS Disclosure; page 239-240; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
XX progression of benign prostatic hyperplasia (BPH), or screening (II) for
XX or identifying an agent that modulates the onset or progression of BPH.
XX The method is based on changes in gene expression in BPH tissue isolated
XX from patients exhibiting different clinical states of prostate
XX hyperplasia as compared to normal prostate tissue. (I) comprises
XX detecting the expression levels of one or more genes in prostate cells
XX from the subject that are differentially regulated compared to normal
XX prostate cells. (II) comprises preparing a first gene expression profile
XX of BPH cells or BPH-like cell population, exposing the cells to the
XX agent, preparing a second gene expression profile of the agent exposed
XX cells, and comparing the first and second gene expression profiles.
XX (I) is useful for diagnosing the onset or progression of BPH. (II) is

PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI: 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1: SEQ ID 3787; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 CC
 XX
 XX

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 9.24e-111 Length: 1157
 Score: 1564.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.92% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-14 (1-317) x ABL65450 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGGTTCTGTGGCTGCTTCTGTCACTTCTGGCAGAGATCCAGGCCAAGGTG 121
 QY 21 GlnGlnAlaValGlnTrpGlnProGlnProGlnLeuArgGlnGlnTrpGlnSer 40
 Db 122 GAGCAACGCGTGGAGACAGACCGGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 181
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 182 GGCACAGCGGTGGAACTGGCACTGGTCTTTGGGATTACCTGGCTGGTGCACACA 241
 QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 242 CTGCTGACGACGAGTGGCAGGAGACTGCTCACTCCCAAGTCAACCAAGTGAAGGCG 301
 QY 81 LeuMetAspGlnTrpMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 302 CTGATGACGACGACCATGAGAGACTGAGAGCTACAAATCGGAATCGAAGAACTGACTG 361
 QY 101 ThrProValAlaGlnGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 Db 362 ACCCGGTAGCGAGGAGACGCGGCGACGGCTGTCCAAAGAGAGTGCAGACGCGCAGGCC 421

QY 121 ArgLeuGlnValaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 422 CGCGTGGCGCCGACATGAGAGAGACGTGGCGCCGCTGTGACATGCCGCGGAGAGTGG 481
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCCATGCTGCGCCAGAGACACCGAGAGACTGGGGTGGCTCGCCCTCCACACTCGCG 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 542 AAGCTGGCTAAGCGGCTCTCCGCGATCCGATACCTGAGAGACCGCTGGCAGTGTAC 601
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaLeuArgGlnArgLeuGly 200
 Db 602 CAGCGCGGGCGCCCGAGGGCGCCGAGCGGCTCAAGCCCATCCCGCAGAGCGCTGGG 661
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220
 Db 662 CCCCTGGTGACAGGCGCGCTGCGGGCGCCACTGTGGGCTCCCTGGCGCGGACAGCCG 721
 QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnMetGly 240
 Db 722 CTACAGGAGGCGGCGCCAGGCTGGGGAGAGCGCTGCGCGCGGATGAGAGATGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
 Db 782 AGTGGACCCGCGACCCCTGGAGAGGTCAAGAGCAGGTGGCGGAGTGGCGGCTCAAG 841
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGAGAGAGCAGGCGCCAGCAGATACGCTTCAGGCGAGGCGCTTCAGGCGCGCTCAAG 901
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlnLeuValGlnLys 300
 Db 902 AGCTGGTTCAGGCGCCCTGGTGGAAAGACATGACAGGCGCAGTGGCGGCTGTGTGAGAG 961
 QY 301 ValGlnAlaAlaValGlnGlnThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCGCTGGGACACAGCCGCGCCCTGTGCGCCAGGCANATTCAC 1012

Search completed: March 14, 2003, 12:42:52
 Job time: 184.728 secs


```

; NAME/KEY: CDS
; LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 4,98e-101
Score: 1146.00
Percent Similarity: 84.06%
Best Local Similarity: 70.31%
Query Match: 72.49%
DB: 4
      Matches: 1126
      Conservative: 225
      Mismatches: 44
      Indels: 43
      Gaps: 3

US-09-827-854-14 (1-317) x US-08-949-155-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 51 ATGAGGGTTCTGTGGTCTTGTGGTACCTCTCTCGCAGATGCCGACAGAGAC 110

QY 21 GluGlnAlaValAluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 111 GAGCCGGGG-----CCGCCCGCGAGGTGCACGTGTGTGGAGGAGGCCCAAG 158

QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 159 TGGCAGGCGCAGCCAGCCCTGGAGCAGGCCCTGGCCCTCTGGGATTACCTGGCCTGG 218

QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 219 GTGCAGTCCCTGTCTACCAAGTGCAGGAGGAGCTCTAGACCAAGGTCCACCCAGGAA 278

QY 78 LeuArgAlaLeuMetAspLysThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 279 CTGACGCGACTGTATGAGAGAGACATGAGAGAGTCAAGCCTACCCGACGAGCTGGAG 338

QY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 339 GCGCAGCTGGGCCCTGACCCAGCAGGAGCGCGCCCTGTCCAAAGAGCTGCAGCGCG 398

QY 118 AlaGlnAlaArgLeuGlnLysAlaAspMetGluAspValArgGlyLeuSerAlaLeuArg 137
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 399 GCGCAGGCGCGGTGGGCGCCGACATGAGAGAGCTGCGACCCGCTGTGCTTACCTCC 458

QY 138 GlyGluValGlnAlaMetLeuGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 459 ACGGAGGTGCACACATGTGGGCGCAGCACCGAGAGACTCGGAGCCGCTGGCTTCC 518

QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 519 CACCTGCGCAAGCTGCGCAAGCGCGCTGCCGACACCGAGGAGCTGCAGAAAGCGCTG 578

QY 178 AlaValTyrGlnAlaGlyAlaArgGluAlaArgGlyLysSerAlaLeuArgGlu 197
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 579 GCGGTGTACAGGCGGGGTGGCGGAGGCGCGAGCGGTGAGGCCCTCCCGCGAG 638

QY 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 639 CCGCTGCGGCGCTGTGTGAGAGAGGCCCATGTCGCGCGCCGACCTGTAGTACAGAGGCC 698

QY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 237
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 699 GCGCAGCCCGCTGCGGAGCGCGGAAAGCTGGGCGCAGAACCTGCGCGAGCGCTGGAG 758

QY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 257
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 759 GAGATGGGAGCGCGGACCGCGACCGCTGATGATGATCGTACAGCTGAGAGCGGTG 818

QY 258 ArgAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 819 CCGACCAAAAGTGGAGAGAGAGGAGCCAGCTTGGCCCTCAGAGCCGAGGAGATTCCAGCGC 878

QY 278 ArgLeuLysSerTrpPheGlnProLeuValGluAspMetGlnArgGlnTrpAlaGluLeu 297
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 879 CTCCTCAAGGCTGCTTGAAGCTCTGGTGAAGACATACGCGCCAGTGGCGCGGCGCTG 938

```

```

QY 298 ValGluLysValGlnAlaValAlaGly---ThrSerAlaAlaProValProSerAspAsn 316
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 939 GTGAGAGAGATGACAGTGGCGGTGAGACATACCTCTCTGCGCCAGTGTATAT 998

RESULT 3
US-09-819-964-5
; Sequence 5, Application US/09819964
; Patent No. 6369294
; GENERAL INFORMATION:
; APPLICANT: Piedadhita, Jorge A
; TITLE OF INVENTION: Compositions and Methods for the
; Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/819,964
; FILING DATE: 28-Mar-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/949,155
; FILING DATE: <unknown>
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hahler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..1001
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-819-964-5

Alignment Scores:
Pred. No.: 4,98e-101
Score: 1146.00
Percent Similarity: 84.06%
Best Local Similarity: 70.31%
Query Match: 72.49%
DB: 4
      Matches: 1126
      Conservative: 225
      Mismatches: 43
      Indels: 8
      Gaps: 3

US-09-827-854-14 (1-317) x US-09-819-964-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 51 ATGAGGGTTCTGTGGTCTTGTGGTACCTCTCTCGCAGATGCCGACAGAGAC 110

QY 21 GluGlnAlaValAluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
   |||.....|.....|.....|.....|.....|.....|.....|.....|
Db 111 GAGCCGGGG-----CCGCCCGCGAGGTGCACGTGTGTGGAGGAGGCCCAAG 158

```

QY 38 TTPGlnSerGlycInArGtrpGluLeuAlaLeuGlyArpPheTrpAspTyrLeuArGtrp 57
Db 159 TGGCAGGCGCAGCCAGCCCTGGAGCAGGCGCTTCTGAGTATCTGCGCTGG 218
QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
Db 219 GTGACAGTCCTGTGTGACCAAGTGCAGAGAGACTGCTCAGCACCAAGTCCACCGAGAA 278
QY 78 LeuArGAlaLeuMetAspGluThrMetGlyGluLeuValAlaTyrGlySerGluGlu 97
Db 279 CTGACGAGAGCTGATAGAGGAGACATGAGAGGTAGAGCCCTACCGCAGAGAGCTGAG 338
QY 98 GluGlnLeuThrProValAlaGlnGluThrArGAlaArGLeuSerGlyGluAla 117
Db 339 GCCAGACTGTGGCCCTGTGACCCAGAGACGAGCGCGCTGTCCAAAGAGAGTGCAGGCG 398
QY 118 AlaGlnAlaArGLeuGlyAlaAspMetGluAspValArGAlaArGLeuValGlnTyrArG 137
Db 399 GCGCAGGCGCGCTGGCGCGGACATGAGAGAGCTGCGCAACCGCTGTGTCTACCGC 458
QY 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGluLeuArGValArGLeuAlaSer 157
Db 459 AGGAGAGTGCACAACTGTTGGGCGCAGACACCGAGAGAGCTGGGAGCGCGCTGCTCC 518
QY 158 HisLeuArGlyLeuArGlyArGLeuLeuArGAspAlaAspAspLeuGlnTyrArGLeu 177
Db 519 CACCTGCGCAAGCTGCGCAAGCGCTGCTCCGCGACACCGAGAGACTGCGCAAGCGCGCTG 578
QY 178 AlaValTyrGlnAlaGlyAlaArGLeuGluGlnAlaGlnArGLeuSerAlaThrArGLeu 197
Db 579 GCCGTGTACCAAGCGCGGCGCTGCGCGAGGCGCGCGAGCGCGCTGAGCGCGCTGCGGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlyArGValArGAlaAlaThrValGlySerLeuAla 217
Db 639 CGCGCTGGGCGCGCTGGGAGAGAGCGCGCTGCGCGCGCGCGCGCGCTGAGTACAGAGGCG 698
QY 218 GlyGlnProLeuGlnGluArGAlaGlnAlaTyrGlyGluArGLeuArGAlaArGMetGlu 237
Db 699 GCGCAGCGCGCTGCGCGAGCGCGCGGAGCGCTGCGCGCGCGCGCGCGCGCGCTGAG 758
QY 238 GluMetGlySerArGTrpArGAspArGLeuAspGluValGlyGlnGluValAlaGlnVal 257
Db 759 GAGATGGCGCAGCGCGCGCGCGCGCGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 818
QY 258 ArgAlaLeuSerGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
Db 819 CGCACAAGAGTGCAGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 878
QY 278 ArgLeuLeuSerTrpPheGluProLeuValGluAspMetGlnArGTrpAlaGlyLeu 297
Db 879 CTGCTCAAAAGGCTGTGCGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 938
QY 298 ValGlnThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
Db 939 GTGAGAGAGTGCAGTGTGCG 998

RESULT 4
US-08-949-155-51
; Sequence 51, Application US/08949155
; Patent No. 6271436
GENERAL INFORMATION:
; APPLICANT: Pledralita, Jorge A
; APPLICANT: Bazer, Fuller W
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ. ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

Alignment Scores:
Pred. No.: 3,14e-87 Length: 4267
Score: 1010.00 Matches: 219
Percent Similarity: 58.65% Conservative: 42
Best Local Similarity: 49.21% Mismatches: 45
Query Match: 63.88% Indels: 140
Gaps: 5

US-09-827-854-14 (1-317) x US-08-949-155-51 (1-4267)

QY 1 MetGlyValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaLeuVal 20
Db 2448 ATGCGTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
QY 21 GlnGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 37
Db 2490 GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2537
QY 38 TTPGlnSerGlyGlnArGtrpGluLeuAlaLeuGlyArGTrpAspTyrLeuArGtrp 57
Db 2538 TGGCAGGCGCAGCCAGCCCTGGAGCAGGCGCGCTGCGCGCTTCTGAGTATCTGCGCTGG 2597
QY 58 ValGlnThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
Db 2598 GTGACAGTCCTGTGTGACCAAGTGCAGAGAGAGACTGCTCAGCACCAAGTCCACCGAGAA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGCACACCGAGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGGGAACCGTGTGTTCTGAGCCCTCAGCTCCAGCCGTCGGGTTCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTGTGCCAACCTCTGGGGGTCTGGGTCTGTCTTTCTTTTCTTCTCTCTCTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTCC 2896
QY 78 ----- 78


```

Db 2837 TTGGGGGAAAAAATTTTCTTTTCTTTCTTCAATTGACTTCAATGCTTCTTCTTCC 2896
Qy 78 ----- 78
Db 2897 ATCTGAGCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
Qy 78 ----- 78
Db 2957 AATCTGTGACGCTCTGCGCATGCGCAGTCAAGAGCCCTCTCTCCCTCACCAGCC 3016
Qy 79 ----- 79
Db 3017 CCCGCCCTCTGCGCCAGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3076
Qy 93 LysSerGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 112
Db 3077 CGCGAGAGAGCTGAGCGAGCGAGCTGCGGCCCGCTGACCCAGAGAGAGAGAGAGAGAG 3136
Qy 113 LysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 132
Db 3137 AAGAGAGCTGAGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3196
Qy 133 LeuValGluThrArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 152
Db 3197 TTGCTCTCTACCGCAGCGAGGTGCAACATTTGGGCGCAGACACCGAGAGAGAGAG 3256
Qy 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArgLysArgLysArg 172
Db 3257 AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3316
Qy 173 LeuGluLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluGluGluGluGlu 192
Db 3317 CTGCAAGACGCGCTGCGCTGTACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3376
Qy 193 SerAlaIleArgGluArgGluGluGluGluGluGluGluGluGluGluGluGluGlu 212
Db 3377 AGCGGCTGCGCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3436
Qy 213 ValGlySerLeuAlaGlyGluProLeuGluGluGluGluGluGluGluGluGluGlu 232
Db 3437 CTAGAGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3496
Qy 233 ArgAlaArgMetGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 252
Db 3497 CGGAGAGCGCTGAGAGAGATGGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3556
Qy 253 GlnValAlaGluValArgAlaLysLeuGluGluGluGluGluGluGluGluGluGlu 272
Db 3557 CAGCTGAGAGAGAGTGGCGCACCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3616
Qy 273 GlnAlaPheGlnAlaArgLeuLysSerTyrPheGluProLeuValGluLysMetGln 292
Db 3617 GAGGATTCACCGCTCTCAAAAGCTGCTGAGACCTGTGTGAAGACATAGCGGCG 3676
Qy 293 GlnTyrAlaGlyLeuValGluGluGluGluGluGluGluGluGluGluGluGluGlu 311
Db 3677 CAGTGGCGCGGCTGTGTGAAGAGATGATGCGCGCTGTGATAGATCTCTCCACCT 3736
Qy 312 ValProSerAspAsn 316
Db 3737 GCGCCCGATGATAT 3751

```

```

ADDRESS: Banner & Wilcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-726-306A-28
Alignment Scores:
Pred. No.: 1,11e-84 Length: 660
Score: 972.00 Matches: 196
Percent Similarity: 98.49% Conservative: 3
Best Local Similarity: 98.49% Mismatches: 0
Query Match: 61.48% Indels: 0
Gaps: 0
US-09-827-854-14 (1-317) x US-08-726-306A-28 (1-660)
Qy 1 MetLysValLeuThrProAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAAGCTTCTGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrPheGln 40
Db 122 GAGCAAGCGGTGGAGAGACAGAGCGGAGCGGAGCTGGCGCAGAGAGAGAGAGAGAG 181
Qy 41 GlyGlnArgTyrPheGluLeuAlaLeuGluGluGluGluGluGluGluGluGluGluGlu 60
Db 182 GCGCAGCGCTGGAGACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGAGCAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysValTyrLysSerGluLeuGluGluGluGlu 100
Db 302 CTGATGAGACGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Qy 121 ArgLeuGlyAlaAspMetGluLysValAlaArgGlyArgLeuValGlnTyrArgGlyGlu 140
Db 422 CGGCTGGCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481

```

```

RESULT 6
US-08-726-306A-28
Sequence 28, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:

```


QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGCTCCAGAGAGCTGCAGGGCGCCAGCCCGGTGGCGCCGACATGAGAGAC 61
QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlnInsThr 148
DB 62 GTGTG-CGCCGCTGTGTGCAGTACCGCGGAGGTGCAGCCATGCTCGCGCAGACACCC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGTGGCTGCCTCCACCTGCACAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAAGTCCCTGCAGCTGACAGCGCGGGCGCCGAGAGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252
RESULT 9
US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-141-24
Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81

Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0
US-09-827-854-14 (1-317) x US-09-287-141-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGCTCCAGAGAGCTGCAGGGCGCCAGCCCGGTGGCGCCGACATGAGAGAC 61
QY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlnInsThr 148
DB 62 GTGTG-CGCCGCTGTGTGCAGTACCGCGGAGGTGCAGCCATGCTCGCGCAGACACCC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGGGGTGGCTGCCTCCACCTGCACAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAAGTCCCTGCAGCTGACAGCGCGGGCGCCGAGAGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252
RESULT 10
US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
Gaps: 0

US-09-827-854-14 (1-317) x US-09-431-613-24 (1-252)

Oy 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGACTGCAGCGCGCGCGCTGGGGCGGACATGAGAGAC 61
Oy 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTG-CGCCCGCTGTGCAGTACCGCGCGAGGTGCAGGCGCATGCTCGGCAGAGCAC 120
Oy 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGCGGGTGCCTCGCCCTCCACCTCGCGCAAGCTGCGTAAGCGCTCTCCGC 180
Oy 169 AspaLaaspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAACTCCCTGCACTGTACAGCGCGGGGCGCGAGGCGGCC 240
Oy 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCGCTC 252

RESULT 11

US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCaulliffe LLP
STREET: 4250 Executive Square, 7th floor
CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037-9103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/504,245

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
Gaps: 0

US-09-827-854-14 (1-317) x US-09-504-245-24 (1-252)

Oy 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGACTGCAGCGCGCGCGCTGGGGCGGACATGAGAGAC 61
Oy 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTG-CGCCCGCTGTGCAGTACCGCGCGAGGTGCAGGCGCATGCTCGGCAGAGCAC 120
Oy 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGCGGGTGCCTCGCCCTCCACCTCGCGCAAGCTGCGTAAGCGCTCTCCGC 180
Oy 169 AspaLaaspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAACTCCCTGCACTGTACAGCGCGGGGCGCGAGGCGGCC 240
Oy 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCGCTC 252

RESULT 12

US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478

GENERAL INFORMATION:

APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th floor
CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037-9103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/287,682

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-397-766-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCACAGAGACTGCGAGCGGCCGCGCTGGGCGGACATGGAGGAC 61
OY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCG-CGCCGCTGTCAGTACCGGGCGAGGTGCAGGCGCATGCTGGCCACAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCGCATGACCTGCGAGAGTCCCTGGCAGTGTACAGGCGGGGCCCGGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 15
US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 6,32e-27 Length: 252
Score: 363.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 22.96% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-14 (1-317) x US-09-287-681-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCACAGAGACTGCGAGCGGCCGCGCTGGGCGGACATGGAGGAC 61
OY 129 ValArgGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCG-CGCCGCTGTCAGTACCGGGCGAGGTGCAGGCGCATGCTGGCCACAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCGCAAGCTGCGTAAGCGGCTCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCGCATGACCTGCGAGAGTCCCTGGCAGTGTACAGGCGGGGCCCGGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

Search completed: March 14, 2003, 20:18:22
Job time : 40.7343 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581

Sequence: 1 MKVLMALVTLFLAGCOAKV.....VEKVOAAGTSAPVPSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp
-O=/cgn2_1/USPFO.spool/US09827854/r/unat_11032003_101612_27578/app_query.fasta.1.3576
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09827854.ecgn_1_1.298.e/unat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=130 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCITUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1581	100.0	1156	10	US-09-827-854-7
2	1573	99.5	1156	9	US-09-870-759-129
3	1573	99.5	1156	9	US-09-802-640-17
4	1573	99.5	1156	10	US-09-827-854-8

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
US-09-827-854-7	1573	99.5	1291	12	US-10-044-090-454	Sequence 454, App																																			
Sequence 7, Appl	1569	99.2	1156	10	US-09-827-854-12	Sequence 12, Appl																																			
Sequence 9, Appl	1565	99.0	1156	10	US-09-827-854-9	Sequence 11, Appl																																			
Sequence 11, Appl	1565	99.0	1156	10	US-09-827-854-11	Sequence 760, App																																			
Sequence 760, App	1564	98.9	1157	10	US-09-954-456-760	Sequence 2244, App																																			
Sequence 2244, App	1564	98.9	1157	10	US-09-880-107-2244	Sequence 5, Appl																																			
Sequence 5, Appl	1558	98.5	1156	10	US-09-827-854-10	Sequence 133, App																																			
Sequence 133, App	1370.5	86.7	41907	10	US-09-967-1013-5	Sequence 374, App																																			
Sequence 374, App	867.5	54.9	786	10	US-09-925-902-133	Sequence 2491, App																																			
Sequence 2491, App	625	39.5	478	10	US-09-964-824-374	Sequence 5420, App																																			
Sequence 5420, App	625	39.5	478	10	US-09-880-107-2491	Sequence 4237, App																																			
Sequence 4237, App	471	29.8	356	10	US-09-960-352-4726	Sequence 9395, App																																			
Sequence 9395, App	466.5	29.5	449	10	US-09-960-352-4726	Sequence 8720, App																																			
Sequence 8720, App	443	28.0	414	10	US-09-960-352-4237	Sequence 14047, A																																			
Sequence 14047, A	439.5	27.8	442	10	US-09-960-352-9395	Sequence 5148, App																																			
Sequence 5148, App	428.5	27.1	416	10	US-09-960-352-8720	Sequence 9797, App																																			
Sequence 9797, App	422.5	26.7	423	10	US-09-960-352-14047	Sequence 130, App																																			
Sequence 130, App	405.5	25.6	409	10	US-09-960-352-5148	Sequence 3497, App																																			
Sequence 3497, App	401.5	25.4	416	10	US-09-960-352-9797	Sequence 1278, App																																			
Sequence 1278, App	389	24.6	253	10	US-09-179-5368-130	Sequence 10599, A																																			
Sequence 10599, A	387.5	24.5	425	10	US-09-960-352-3497	Sequence 4187, App																																			
Sequence 4187, App	380.5	24.1	390	10	US-09-960-352-1311	Sequence 8042, App																																			
Sequence 8042, App	379.5	24.0	391	10	US-09-960-352-1278	Sequence 3235, App																																			
Sequence 3235, App	379.5	24.0	392	10	US-09-960-352-10599	Sequence 14063, A																																			
Sequence 14063, A	379.5	24.0	401	10	US-09-960-352-4187	Sequence 24, Appl																																			
Sequence 24, Appl	375.5	23.8	413	10	US-09-960-352-8042	Sequence 176, App																																			
Sequence 176, App	375.5	23.4	404	10	US-09-960-352-2325	Sequence 4914, App																																			
Sequence 4914, App	369.5	23.4	405	10	US-09-960-352-14063	Sequence 10040, A																																			
Sequence 10040, A	363	23.0	252	10	US-09-796-416-24	Sequence 10361, A																																			
Sequence 10361, A	363	23.0	252	10	US-09-879-341-24	Sequence 11986, A																																			
Sequence 11986, A	351.5	22.2	377	10	US-09-960-352-176	Sequence 3457, App																																			
Sequence 3457, App	350.5	22.2	353	10	US-09-960-352-4914	Sequence 10540, A																																			
Sequence 10540, A	350.5	22.2	378	10	US-09-960-352-10040	Sequence 8934, App																																			
Sequence 8934, App	350.5	22.2	386	10	US-09-960-352-10361	Sequence 6552, App																																			
Sequence 6552, App	348.5	22.0	388	10	US-09-960-352-11986	Sequence 5408, App																																			
Sequence 5408, App	344.5	21.9	396	10	US-09-960-352-11986	Sequence 3681, App																																			
Sequence 3681, App	344.5	21.8	377	10	US-09-960-352-10540																																				
	338.5	21.4	432	10	US-09-960-352-8934																																				
	338	21.4	432	10	US-09-960-352-6552																																				
	337.5	21.3	376	10	US-09-960-352-5408																																				
	335.5	21.2	373	10	US-09-960-352-3681																																				

ALIGNMENTS

RESULT 1
US-09-827-854-7
Sequence 7, Application US/09827854
Patent No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
APPLICANT: Kyriacs, Kyriacos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-7
Alignment Scores:
Pred. No.: 7.4e-134
Score: 1581.00
Percent Similarity: 100.00%
Length: 1156
Matches: 317
Conservative: 0

Db 661 CCCCTGGTGAACAGAGGCCCGTGGGCGCCGCTGAGCTCCCTGGCCGCCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCGCTGGGCGCGAGCGGCTGGCGCGCGGATGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 ACCCGAGACCGCGGACCGCTGGACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAG 900
QY 281 SerTrpPheGluProLeuValGluLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 ACCTGCTTCAGACCCCTGGTGGAAACATGACAGCCGAGTGGGCGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGCTGGGACACAGCCCGCCCTGTGCCAGGACATAC 1011
RESULT 3
US-09-802-640-17
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsel Aruna
; APPLICANT: Kleya Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802, 640
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1014)
; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
; OTHER INFORMATION: (APOE)
US-09-802-640-17
Alignment Scores:
Pred. No.: 3,88e-133 Length: 1156
Score: 1573.00 Matches: 316
Percent Similarity: 99.68% Conservatave: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.49% Indels: 0
Gaps: 0
US-09-827-854-14 (1-317) x US-09-802-640-17 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTCTCTGGGCTGGCGCTGCTGTCATCTCGGAGATGCCAGGCCAAGG 120
QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCGGAGCGCGGAGCGGAGAGAGAGAGAGAGAGAGAG 180
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyLeuArgTrpValGlnThr 60
Db 181 GGCACAGCGGTGGAGACTGGGCTGCTTTGGGATTAACCTCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuValArgAla 80
Db 241 CTGTCTGAGCAGGTGACAG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGAGAGACCATGAGAGAGTTGAAGGCTTCAATTCGAACTGGAGAGAACTG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGGGAGAGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTrpArgGlyGlyVal 140
Db 421 CGGCTGGGCGGCGGACATGAGAGAGCTGTGGCGCCCTGTGGCTACATACCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyLysSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCAATGCTGGGCGAG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
Db 541 AAGCTGGGTAAAGCGGCTCTCCGCGATGCCATGACCTGAGAAAGCGCTGGCACTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 601 CAGCGCGGCGGCGCGGAGAGGCGCGAGCGCGCTCACGCGCATCCGAGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
Db 661 CCCCTGGTGAACAGAGCGCGCTGGGCGCGCTGAGCTGGGCTCCCTGGCGGACAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaThrPgluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGACCGCGGACCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAG 900
QY 281 SerTrpPheGluProLeuValGluLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGCTTCAGACCCCTGGTGGAAACATGACAGCCGAGTGGGCGGCGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGCTGGGACACAGCCCGCCCTGTGCCAGGACATAC 1011
RESULT 4
US-09-827-854-8
; Sequence 8, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Kyriacos, Kyriakos E.
; APPLICANT: Zannis, Vassilis
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827, 854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679, 088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544, 386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-8

Db 726 CAGCCCGGGCCCGGAGGGCCGCGGCTCAGCGCCATCCGAGCGCCCTGGG 785
Qy 201 ProleuValGluGluGluValArgValAlaAlaThrValGlySerLeuAlaGlyGluPro 220
Db 786 CCCCTGCTGGAACAGAGCGGCGTGGGGCGCCACTGTGGCTCCCTGGCCGCAAGCCG 845
Qy 221 LeuGluGluArgAlaGluAlaThrPgluValArgLeuArgAlaArgMetGluGluMetGly 240
Db 846 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGCGATGAGAGATGGGCG 905
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGluGluValAlaGluValArgAlaLys 260
Db 906 ACCCGAGACCGCGGACCGCTGGTGGACGAGTGAAGAGAGAGTGGCGAGTGGCGCCCAAG 965
Qy 261 LeuGluGluGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 280
Db 966 CTGGAGAGAGAGCGCCAGCATACGCTGGAGCGCGAGGCTTCACAGGCGCCGCTCAAG 1025
Qy 281 SerTrpPheGluProLeuValGluAspMetGluArgGlnTrpAlaGlyLeuValGluLys 300
Db 1026 ACTGTGTTGAGACCCCTGGTGGAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 1085
Qy 301 ValGluAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 1086 GTGCAGGCTGCGCGTGGGACACGACGCGCCCTGTGCCAGCAGACATCAC 1136
RESULT 6
US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 2000-10-04
; PRIOR FILING DATE: 2000-10-04
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-12
Alignment Scores:
Pred. No.: 8,89e-133 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-14 (1-317) x US-09-827-854-12 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGluAlaLysVal 20
Db 61 ATGAGAGTTCTGTGGCTGGCTGCTGTCATCTCTGGCGAGATGCGACGCAAGGTG 120
Qy 21 GluGluAlaValAlaGluThrGluProGluProGluLeuArgGluGluGluGluGluGluGlu 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGCGCCAGCAGACAGAGTGGCGAGAGC 180
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGluValArgPheTrpAspTyrLeuArgTrpValGluThr 60
Db 181 GGGCAGCGCTGGAGACATGCGACTGGCTTTTGGGATTCACCTGCGTGGTGCAGACA 240
Qy 61 LeuSerGluGluValGluGluGluLeuSerSerGluValThrGluGluLeuArgAla 80

Db 241 CTGCTGAGACAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCACCAGAACTGAGGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
Db 301 CTGATGAGACAGAACCATGATGAGAGTGAAGGCCCTCAATCGGAACCTGGAGAACAACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaGluAla 120
Db 361 ACCCGGTGGCGGAGAGAGACGGGAGCGGCTGTCCAAAGAGACTTCAGAGCGGCGAGGCC 420
Qy 121 ArgLeuGluAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyLysVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCCGCCCTGTGTGCACTACCGCGGCGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGCAAGACACGAGAGAGCTGGCGGTGGCGCTCCCTCCACCTGGCG 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
Db 541 AAGTGCCTCAGCGGCTCTCCGCGATGCCATGACCTGCAAGAGCGCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGCGCGGAGGGCGCGAGCGGCGCTCAGCGCATTCGCGAGCGCTGGGG 660
Qy 201 ProleuValGluGluGluValArgValAlaAlaThrValGlySerLeuAlaGlyGluPro 220
Db 661 CCCCTGTGGGAACAGGGCGGCTGGGGCGCGCCACTGTGTGGCTCCCTGGCGGCGAGCGG 720
Qy 221 LeuGluGluArgAlaGluAlaThrPgluValArgLeuValArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGluValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCCCGGAGCGCGCTGGAGCAGGTGAAGAGAGAGTGGCGGAGTGGCGCCCAAG 840
Qy 261 LeuGluGluGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 280
Db 841 CTGGAGAGAGAGCGCCAGCATACGCTGCAGGCGCGAGGCTTCACAGCGCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGluArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTGTCAGACCCCTGTGGTGAACATGTCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
Qy 301 ValGluAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCCAGGAGACATCAC 1011
RESULT 7
US-09-827-854-9
; Sequence 9, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT FILING DATE: 2001-04-05
; PRIOR FILING DATE: 2000-10-04
; PRIOR FILING DATE: 2000-10-04
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-827-854-9

Alignment Scores:

Pred. No.:	2,04e-132	Length:	1156
Score:	1565.00	Matches:	315
Best Local Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	98.99%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-14 (1-317) x US-09-827-854-9 (1-1156)

```
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB ATGAAGTTCTGTGGCTGGTGTGCTGTGCATTCCTCGCAGAGATCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluProGln 40
DB 121 GACCAAGCGGTGAGACAGAGCCGAGCCGACCTGCGCAGACGACCGAGTGGCAGAGC 180
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCAGCGCTGGGAGACTGGGACCTGGCTTTGGGATTACTGGCTGGTGGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGAGCAGCTGCAGAGAGCTGTCTCAGCTCCAGAGTACCAGAGAGTGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 301 CTATGAGCAGACCAATGAAGAGTTGAAGCCTTACAAATCGGAGCTGGAGAACATG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGCTGGCGAGAGAGCGGGCAGCGCTGTCCAGAGAGCTGCAGCGGGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGCGACATGTGGAGCTGTGGCGGCTGTGTGACATCCCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCACAGACACCGAGAGACTCGGGTGGCTCGCTCCACCTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTCGTAAGCGGCTCCTCCGCGATGCCGATGACCTGCAGAAAGTCCGCGAGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluArgGlyArgLeuSerAlaIleArgGluLeuArgLeuGly 200
DB 601 CAGCGCGGGCGCCGAGCGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGTGGAACAGGCGCGCTGGCGGCCGCCACTGTGGGCTCCTGGCGCGCCAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 721 CTACAGAGAGGGGCCAGGCGCTGGCGGAGCGGCTCGCGCGCGGATGGAGAGTGGGG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 ACCCGGACCCCGGACCGCTGTGACGAGGTGAAGACAGGAGTGGCGAGGTGGCGCGCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 841 CTGGAGAGACAGGCCCAAGAGATACGCTGTGACAGGCCGAGGCCCTTCCAGGCCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGTGTGAGGCCCTGTGTGGAAGACATGACAGCGCCAGTGGCGGTGTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
```

```
DB 961 GTGACAGCTGCGTGGGACACAGCGCGCCCTGTGCTCCAGCAGACATATCAC 1011
```

RESULT 8

```
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-11
```

Alignment Scores:

Pred. No.:	2,04e-132	Length:	1156
Score:	1565.00	Matches:	315
Percent Similarity:	99.37%	Conservative:	0
Best Local Similarity:	99.37%	Mismatches:	2
Query Match:	98.99%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-14 (1-317) x US-09-827-854-11 (1-1156)

```
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGCTGGTGTGCTGTGCATTCCTCGCAGAGATCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluProGln 40
DB 121 GACCAAGCGGTGAGACAGAGCCGAGCCGACCTGCGCAGACGACCGAGTGGCAGAGC 180
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCAGCGCTGGGAGACTGGGACCTGGCTTTGGGATTACTGGCTGGTGGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGAGCAGCTGCAGAGAGCTGTCTCAGCTCCAGAGTACCAGAGAGTGGGCGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 301 CTATGAGCAGACCAATGAAGAGTTGAAGCCTTACAAATCGGAGTGGAGAACATG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGCTGGCGAGAGAGCGGGCAGCGCTGTCTCAGAGAGCTGCAGCGGCCGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValArgGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 421 CGGCTGGGCGCGACATGTGAAGACGTGTGGCGGCTGTGTGACATCCGAGTACCGCGCGAGTGG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCACAGACACCGAGAGCTGGCGGTGGCTCGCTCCACCTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGTGTAAGCGGCTCCTCCGCGATGCCGATGACCTGCAGAAAGTCCGCGAGTAC 600
```


QY	181	GLIAGAGIAlAArvgIuagIuAlaGIArGlyIuSerAlAlIeArGluArGluGly	200
Db	601	CAGCCGGGGCCCGCGAGGGCGCCGAGCGGGCTCAAGCCATCCGAGCGCCCTGGGG	660
QY	201	ProLeuValGIuGIuGIuArGValArGAlAlAlArThrValISerLeuAlaGIuGIuPro	220
Db	661	CCCTGGTGGAAcAGGGGCCCGTGGGGGCCCACTGGGCTTCCTGGCGGCCAGCCG	720
QY	221	LeuGIuGIuArGAlaGIuAlAlArTPGlyGIuArGIuArGAlaArGmetGIuIuMetGIy	240
Db	721	CTACAGAGAGGGGGCCAGGCCCTGGGGGAGACGGCTGGCGCGGATGAGAGAGATGGGC	780
QY	241	SerArGThrArGAspArGleuAspGIuValLyGIuGIuAlaGIuAlaGIuAlaGIuAlaGly	260
Db	781	AGCCGAGCCCGCGACCCCTGTGAGAGAGTAAAGAGAGAGGTGGCGGAGGGCCGCCGAAG	840
QY	261	LeuGIuGIuGIuAlaGIuGIuAlIeArGIuGIuAlaGIuAlaGIuAlaGIuAlaArGleuGly	280
Db	841	CTGAGAGAGCAGGCCCGACGCAGATACGCTCTCAGGCCGAGGCCCTTCCAGAGGCCCGCTCAAG	900
QY	281	SerTPheGIuProLeuValGIuAspMetGIuArGIuAlaGIuAlaGIuAlaGIuAlaGly	300
Db	901	AGCTGATTCCAGGCCCTGTGTGAAGACATCATCACGCCAGCGAGGGCGGTGTGTGAGAG	960
QY	301	ValGIuAlaAlaValGIyThrSerAlaAlaProValProSerAspAsnHis	317
Db	961	GTTCAGAGCTGCGGTGGGCACAGGCCCGCCCTGTGCCACGCACATATCAC	1011

RESULT 9	
US-09-954-456--760	
Sequence 760, Application US/09954456	
Patent No. US20020115057A1	
GENERAL INFORMATION:	
APPLICANT: Young, Paul	
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can	
TITLE OF INVENTION: Sets	
FILE REFERENCE: 689290-76	
CURRENT APPLICATION NUMBER: US/09/954,456	
CURRENT FILING DATE: 2001-09-18	
PRIOR APPLICATION NUMBER: US/60/233,617	
PRIOR FILING DATE: 2000-09-18	
PRIOR APPLICATION NUMBER: US/60/234,052	
PRIOR FILING DATE: 2000-09-20	
PRIOR APPLICATION NUMBER: US/60/234,923	
PRIOR FILING DATE: 2000-09-25	
PRIOR APPLICATION NUMBER: US/60/235,134	
PRIOR FILING DATE: 2000-09-25	
PRIOR APPLICATION NUMBER: US/60/235,637	
PRIOR FILING DATE: 2000-09-26	
PRIOR APPLICATION NUMBER: US/60/235,638	
PRIOR FILING DATE: 2000-09-26	
PRIOR APPLICATION NUMBER: US/60/235,711	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: US/60/235,720	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: US/60/235,840	
PRIOR FILING DATE: 2000-09-27	
PRIOR APPLICATION NUMBER: US/60/235,863	
PRIOR FILING DATE: 2000-09-27	
NUMBER OF SEQ ID NOS: 2276	
SOFTWARE: PatentIn version 3.0	
SEQ ID NO 760	
LENGTH: 1157	
: TYPE: DNA	
: ORGANISM: Homo sapiens	
US-09-954-456--760	
Alignment Scores:	
Pred. No.:	2,51e-132
Score:	1564.00
Percent Similarity:	99.05%
Best Local Similarity:	99.05%
Query Match:	98.92%
	Length: 1157
	Matches: 314
	Conservative: 0
	Mismatches: 3
	Indels: 0

DB:	10	Gaps:	0
US-09-827-854-14 (1-317) x US-09-954-456-760 (1-1157)			
QY	1	McttysValLleuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyGlyGlnAlaLeuVal	20
DB	62	ATGAAGGTTCTCTGGGCTGCGCTGGTGCATTCTCGCAGAGATGCCAGCCAAAGTG	121
QY	21	GtGlnAlaValaGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer	40
DB	122	GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCAGCAGACCAAGTGGCAGAGC	181
QY	41	GtGlnAlaGtTPGluLeuAlaLeuGlyAArgPheTrpAspIlyLeuAgtTPValGlnThr	60
DB	182	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTTGGGATTACCTGCCTGGGTCCAGACA	241
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuAArgAla	80
DB	242	CTGTCTGACAGAGTCCAGAGAGAGCTGCTCAAGTCCCAAGTCACCCAAAGATCAGGGCG	301
QY	81	LeuMeLAspGluThrMetIlyGluLeuLeuysAlaTrpIlySerGluLeuGluGlnLeu	100
DB	302	CTGATGGAGCAGAACCATGAGAGAGTGAAGGCTCAACAATCGGAACCTGAGAGAACATG	361
QY	101	ThrProValAlaGluGluThrAArgAlaArgLeuSerIlyGluLeuGlnAlaAlaGlnAla	120
DB	362	ACCCCGGTAGCGGAGAGACGCGGGCACGCGTGTCCAAAGAGCTGCAGACGCGCACAGCC	421
QY	121	ArgLeuGlyAlaAspMetGluAspValAArgIlyTrpLeuValGlnTrpArgGlyGluVal	140
DB	422	CGGCTGGGCGCGACATGAGAGACCTGTGCGGCGCCGCTGGTGCACTACCGGGGAGGTG	481
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160
DB	482	CAGGCCATGCTCGGCCAGAGACCGAGAGAGCTGGGGGGCGCTCGGCTCCACCTGCGC	541
QY	161	LysLeuAArgIlyAArgLeuLeuAArgAspAlaAspAspIleuGlnIlyAArgLeuAlaValTyr	180
DB	542	AAGCTGCTAAGCGGCTCTCTCCGGATCCCATGACCTGCAGAAAGCCCTGGCGAGTGTAC	601
QY	181	GlnAlaGlyAlaAArgGluGlyValaGluAArgIlyLeuSerAlaIleAArgGluAArgLeuGly	200
DB	602	CAGGCCGGGGCGCCGGAAGGGCGCGAGCGCGGCTCAAGCGCCATGCCGAGCGCTGGGG	661
QY	201	ProLeuValaGluGlnGlyAArgValAArgAlaIleAThrValaGlySerLeuAlaGlyLPro	220
DB	662	CCCCGTGGGAACAAGGGCGCGGTGCGGGCCGCCACTGTGGGCTCCCTGCGCGCCACCG	721
QY	221	LeuGlnGluAArgAlaGlnAlaIleAArgIlyGluAArgLeuAArgAlaAArgMetGluMetGly	240
DB	722	CTACAGGAGAGCGGGCCCAAGGCTGGGGGCGAGGGGCTGCGCGCGGAGTGAAGATGGGC	781
QY	241	SerAArgTrpAArgAspAArgLeuAspGluValIlyGlnGlnValaAlaGluValAArgAlaLys	260
DB	782	AGTCGAGCCCGCGAGCCGCTGAGCAGAGGTGAAGAGCAGAGTGGGGAGTGGCGGCAAG	841
QY	261	LeuGlnGluGlnAlaGlnGlnIleAArgPheGlnAlaGlnAlaPheGlnAlaAArgLeuLys	280
DB	842	CTGGAGGAGAGCAGCCCAACAATACGCTGCAGGGCCGAGGCGCTTCCAGGCCCGCTCAAG	901
QY	281	SerTrpPheGluProLeuValaGluAspMetGlnAArgGlnTrpAlaGlyLeuValaGluLys	300
DB	902	AGCTGCTTCGAGCCCTGTGGTGAAGACATGCAGGCGCCAGTGGCGGGGTGTGAGAGAG	961
QY	301	ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis	317
DB	962	GTCGAGGCTGCCGTGGGACACAGGCGCGCCCTGTGGCCAGCGACATCTCAC	1012
RESULT 10			
US-09-880-107-2244			
; Sequence 2244, Application US/09880107			
; Patent No. US20020142981A1			
; GENERAL INFORMATION:			

Oy	158	sleuNrArGylseuAArgYlsARgIeuleNuArgrspblAspsAlseuNgInlvsArgIeul	178
Db	21336	CCTGGCGAAAGCTGYrTAAGCGGCTCTCCGCATCSCCGTACCTGCAGAAAGGCCCTGGC	21395
Oy	178	aValTrgslnAlagIalYAarGslUGlYlAgIuArGrsglyseuSerAlAlearGluAr	198
Db	21396	AGTGnACCAGGCGGGGGCCCCGGAGGGGGCGGAGCGGGCCrCACGGCCATCCGGAGCG	21455
Oy	198	glEuGlYPrOleuVAIlGUngInglyArGyalArAlaalArThrValIGlyserLeuAlGI	218
Db	21456	CCTGGGGCCCCGTGGGGAACAAGGGCGGGGTGGGGCCGCACACTGTGGCTCCCTGGCCGG	21515
Oy	218	yGlnPrlOuEngIngluArGAlaglnAlaETrgpLyGuArGrLeuArGAlArGrMetGlnGl	238
Db	21516	CCAAGCCGCTACAGAGAGCGGGGCCCAAGCCTCGGGGGCCAGCGCGCTGCCCGCATGGAGA	21575
Oy	238	uMeGlYSeArArThrArGAsPArGLEuaSPGLuVallysGlUgnlValAlaglVuAlAr	258
Db	21576	GATGGGACAGCCGGAGCCCCGGACCGCTGGAGCAGAGTGAAGGAGCAGGTGGCGAGGTGGG	21635
Oy	258	gAlAlArSleuGlUGlUGlnAlagInGlnlIeArGrleuGlnAlagluAlaphEGlnAlAr	278
Db	21636	CCCACAAGCTGAGGAGCAGAGCCCAACACAGATACGCTGCAGAGCGGAGGCTTCACAGGCCG	21695
Oy	278	gleuNysSeTrPPheGluProLeuVAlGUAsPMetGlnArGrGlnTrTPAlaGlyLeuVa	298
Db	21696	CCTCAAGAGCTGTTCGAGCCCTCTGGTAGACATGTCAAGCCCAAGTGGCGGGCTGGT	21755
Oy	298	IgluYsValGlnAlaAlaValaIGlYThrSerAlaAlaProValProSerASPasnHis	317
Db	21756	GGAGAAAGGTGCAGCGTGGCGGTGGACACAGCGCCGCCCTGTGGCCAGGACAATCAC	21813
RESULT 13			
US-09-925-302-133			
; Sequence 133, Application US/09925302			
; Patent No. US20020044941A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
; FILE REFERENCE: PA104			
; CURRENT APPLICATION NUMBER: US/09/925.302			
; CURRENT FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: PCT/US00/05918			
; PRIOR FILING DATE: 2000-03-08			
; PRIOR APPLICATION NUMBER: 60/124,270			
; PRIOR FILING DATE: 1999-03-12			
; NUMBER OF SEQ ID NOS: 896			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 133			
; LENGTH: 786			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-302-133			

Alignment Scores:	
Pred. No.:	7.37e-70
Score:	867.50
Percent Similarity:	98.39%
Best Local Similarity:	98.39%
Query Match:	54.87%
DB:	10
	Length: 786
	Matches: 183
	Conservative: 0
	Mismatches: 3
	Indels: 2
	Gaps: 0

Oy	132	ArgLeuValGlnTyrArgGlyGlnValGlnImlaMetLeuGlyGlnSerThrGlnGluLeu	151
Db	6	CGCGTGGCAGTACCGCGGGAGGTGCASCCATGCTCGGCCAGACACCGAGACTG	65
Oy	152	ArgValArgLeuAlaSerHisLeuArgGlyLeuArgGlySerArgLeuArgAspAlaAsp	171
Db	66	CGGGGCGCCCTGGCTCCACACTGCGGAA-CTGGGTAA-CGGCTCTCTCGCGATGCCGAT	123

OY	172	AsLeuGlnLysArGLeuAlaIValTyrGlnAlaIArgLysAlaArgLysAlaArgLys	191
Db	124	GNCTGGCAAGAGCCCTGTGGCACTACAGAGCCGGGGCCCGCAAGGGCCCGAGCCGGCC	183
OY	192	LeuSerAlaIleArGArgLysGluArgLeuGlyProLeuValGluGlnGlyArgValArgAlaIa	211
Db	184	CTCACGGCCATTCGGAGAGCCCTGGGGGCCCTGTGTGGAAACAGGGCCCGTGGGGGGCCGC	243
OY	212	ThrValGlySerLeuAlaGlyGlnProLeuGlnLysArgAlaGlnAlaITPGLysLysArg	231
Db	244	ACTGTGGGCTCTCCGTGGCCGGCCACCCGTACAGAGGGGGGCCACAGGCTTGGGGGAGACGG	303
OY	232	LeuArGAlaArGMeGluGluMeGlySerArGThrArGAspArgLeuAspGluValLys	251
Db	304	CTGGCGCGGGAGTGGAGAGATGGGAGCCGGACCCGGACCGCCCTGTGACGAGGTCAAG	363
OY	252	GluGlnValAlaGluValArGAlaIbLysLeuGluGluGlnAlaGlnGlnIleArGLeuGln	271
Db	364	GAGCAGGTGGGGAGGTGGCCGCAAGCTGAGGAGGAGGCCACGATACGGCTTCAG	423
OY	272	AlaGluAlaPheGlnAlaArGLeuLysSerTyrPheGluProLeuValGluAspMeGln	291
Db	424	GCCAGAGGCTTCCAGAGCCCGGCTCAAGAGCTGTTCAGAGCCCTGTGTGAAAGCATTCAG	483
OY	292	ArgGlnTPPALaGlyLeuValGluLysValGlnAlaIaValAlaGlyThrSerAlaIaPro	311
Db	484	CGCCAGTGGGGCGGGCTGTGTGGAAAGGTGCAAGGCTCCGTGGGCAACAGCGCGCCCT	543
OY	312	ValProSerAspAsnHis	317
Db	544	GTGCCACGACGACATTCAC	561

```

US-09-964-824A-374/c
: Sequence 374. Application US/09964824A
: Patent No. US20020102531A1
:
: GENERAL INFORMATION:
:   APPLICANT: Horrigan, Stephen
:   TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
:   TITLE OF INVENTION: Sets
:   FILE REFERENCE: 689290-73
:   CURRENT APPLICATION NUMBER: US/09/964,824A
:   CURRENT FILING DATE: 2001-09-27
:   PRIOR APPLICATION NUMBER: US/60/236,033
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,032
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,028
:   PRIOR FILING DATE: 2000-09-28
:   NUMBER OF SEQ ID NOS: 583
:   SOFTWARE: PatentIn version 3.0
:   SEQ ID NO 374
:   LENGTH: 478
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:   NAME/KEY: misc_feature
:   LOCATION: (1)...(478)
:   OTHER INFORMATION: n-a,t,g or c
US-09-964-824A-374

```

Alignment Scores:	
Pred. No.:	2,696-48
Score:	625.00
Percent Similarity:	97.69%
Best Local Similarity:	96.92%
Query Match:	39,538
DB:	10
Length:	478
Matches:	1266
Conservative:	1
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-834-14 (1-317) x US-09-964-824A-374 (1-478)

Qy 1 MettysValleuTriplaAlaLeuValPhrPheLeuAlaGlyCysGlnAlaVal 20
|||||

Db 396 ATGAGGTTCTGTGGGCTGCTGTGTCACATTCTTGAGGATGCCAGGCCAAGGTG 337
QY 21 GtGlnAlaValGluThrGluProGluLeuArgGlnGluThrGlnTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACACAGCCGGAGCCNCAGACTGCGCAGAGAGCCAGGTGGCAGAGC 277
QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGTGGGAACGTGCTGCTGCTTTGGATTACTGCGCTGGGTGCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValTrpGlnGluLeuArgAl 80
Db 216 CTGTCTGACAGAGTGCAGAGAGAGTGTCTGAGCTCCAGTCCACCAGCACTGAGGCGC 157
QY 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
Db 156 GGTGATGAGCAGACCATGAAAGAGTGAAGGCTTCAAAATCGAACTGAGAACCACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGGAGAGAGACCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGAGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGCATGAGAGACGTG 9
RESULT 15
US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491
Alignment Scores:
Pred. No.: 2.69e-48
Score: 625.00 Length: 478
Percent Similarity: 97.69% Matches: 126
Best Local Similarity: 96.92% Mismatches: 1
Query Match: 39.53% Indels: 2
DB: 10 Gaps: 0
US-09-827-854-14 (1-317) x US-09-880-107-2491 (1-478)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAGGTTCTGTGGGCTGCTGTGTCACATTCTTGAGGATGCCAGGCCAAGGTG 337
QY 21 GtGlnAlaValGluThrGluProGluLeuArgGlnGluThrGlnTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACACAGCCGGAGCCNCAGACTGCGCAGAGAGCCAGGTGGCAGAGC 277

QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGTGGGAACGTGCTGCTGCTTTGGATTACTGCGCTGGGTGCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValTrpGlnGluLeuArgAl 80
Db 216 CTGTCTGACAGAGTGCAGAGAGAGTGTCTGAGCTCCAGTCCACCAGCACTGAGGCGC 157
QY 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
Db 156 GGTGATGAGCAGACCATGAAAGAGTGAAGGCTTCAAAATCGAACTGAGAACCACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGGAGAGAGACCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGAGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGCATGAGAGACGTG 9
Search completed: March 15, 2003, 03:03:28
Job time : 82.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-14

Perfect score: 1581

Sequence: 1 MKVLMALVTLFLGCAQV.....VEKQAAVGSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n_model -DEV=xdp
-O=/cgn2.1/USPTO.spool/US09827854/runal_11032003_101610_27495/app_query.fasta_1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -NORM=ext -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.cgn.1.1.8826.grunal_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1278.5	80.9	1027	13 B1670367
2	1242	78.6	842	12 B6763371
3	1236	78.2	933	13 B1668318
4	1208	76.4	938	12 B6761746
5	1205.5	76.2	942	13 B1600906
6	1191	75.3	800	13 B6042094
7	1190.5	75.3	927	12 B6472299
8	1187.5	75.1	922	13 B1597743
9	1178.5	74.5	817	12 B6774871
10	1177	74.4	811	13 B1600563
11	1163.5	73.6	845	12 B6829472
12	1140	72.1	790	12 B6707147
13	1137	71.9	919	13 B1551475
14	1136	71.9	706	14 B6728696
15	1134	71.7	1110	11 AK010261
16	1127	71.3	757	13 B6042228
17	1117.5	70.7	907	12 B6706129
18	1116	70.6	741	12 B6762924
19	1112	70.3	804	12 B6702752
20	1111	70.3	834	13 B6042676
21	1108.5	70.1	808	13 B1668329
22	1105	69.9	855	13 B161362
23	1102	69.7	803	13 B1670350
24	1099	69.5	782	12 B6716776
25	1099	69.5	912	13 B1601551
26	1099	69.5	865	14 B6677266
27	1096	69.3	757	13 B1603658
28	1093.5	69.2	797	12 B6715366
29	1093	69.1	846	13 B1159757
30	1092	69.1	796	13 B6042153
31	1092	69.1	1100	14 B6914382
32	1087	68.8	757	12 B6707750
33	1087	68.8	794	13 B1601279
34	1087	68.8	954	12 B6967543
35	1085	68.6	790	13 B1551066
36	1084	68.6	790	13 B1551811
37	1080.5	68.3	798	12 B6708414
38	1079.5	68.3	891	13 B1549232
39	1073.5	67.9	802	13 B1458355
40	1071.5	67.8	914	13 B1603523
41	1071	67.7	748	13 B1553085
42	1070.5	67.7	930	12 B6967857
43	1066.5	67.5	812	12 B6769968
44	1066	67.4	688	13 B1544866
45	1065.5	67.4	803	12 B6709360

ALIGNMENTS

RESULT 1
B1670367.
LOCUS 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
DEFINITION B1670367 1027 bp mRNA linear EST 12-SEP-2001
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS mRNA sequence.
SOURCE B1670367.1
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11790 row: 1 column: 09
 High quality sequence stop: 845.
 Location/Qualifiers
 1. 1027
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5312024"
 /tissue_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site.1: BamHI; Site.2: SalI; xhoI (gtcgag
): Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVA-3',
 size selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 194 a 308 c 400 g 125 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,79e-124 Length: 1027
 Score: 1278.50 Matches: 291
 Percent Similarity: 90.55% Conservative: 6
 Best Local Similarity: 88.72% Mismatches: 20
 Query Match: 80.87% Indels: 12
 DB: 13 Gaps: 2
 US-09-827-854-14 (1-317) x B1670367 (1-1027)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 DB 26 ATGAAGGTTCTGTGGGTGCTGTGCTGCATTCCTGCAGAGATGCCAGGCCAAGGTC 85
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnProGlnSer 40
 |||
 DB 86 GAGCAAGCGGTGAGACAGAGCCGAGCCGACCTGGCCAGACAGACCGAGTGGCAGAGC 145
 QY 41 GlyLysArgTrpLeuLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||
 DB 146 GGGCAGCGCTGGGAGACCTGGGCTGTTGGATTACTGCGCTGGGTGGCAGACA 205
 QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||
 DB 206 CTCTCTGAGCAGCTGCAGGAGGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGGCG 265
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaThrLysSerGlnLeuGlnGlnGln 100
 |||
 DB 266 CTGATGAGCAGACCATGAAGAGGTTGAGGCTTACAAATCGGAGACTGGGAGAACACTG 325
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 |||
 DB 326 ACCCGCGTGGCGGAGAGACGGCGGCGCTTCAAGAGCTGCAGCGGGCGCAGGCC 385
 QY 121 ArgLeuGlyAlaAspMetGlnAspValArgGlyArgLeuValGlnThrArgGlyGlnVal 140
 |||
 DB 386 CGGCTGGGGCGGACATGTGAGGACGTGTGGCGCGCTGTGACATACCGCGCGAGGTG 445
 QY 141 GlnAlaMetLeuGlnGlnThrGlnGlnGlnValArgLeuAlaSerHisLeuArg 160
 |||

DB 446 CAGGCCATGCTGGCCAGACAGACCAGAGAGCTCGGGTGGCTCCACCTGCGCC 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 |||
 DB 506 AACCTGGGTAAAGGGCTCTCCCGGATGCCGATGACCTGACAGAGCCCTGGCAGTTCAC 565
 QY 181 GlnAlaGlyAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeu 200
 |||
 DB 566 CAGCGCGGGCGCCGAGCGGCGCGAGCGCGGCTCAGCCCATCCGCGAGCCGCTGGGG 625
 QY 200 yProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnTr 220
 |||
 DB 626 CCCCTGTGGTGAACAGGTCCGCTGGCGCGCCGCTGCTCCCTGCGCGCCAGCC 665
 QY 220 OleuGlnGlnArgAlaGlnAlaTrpGlnArgGlyArgGlyLeuArgAlaArgMetGln 239
 |||
 DB 686 GCTACAGAGACGGGCCAAGGCTGGGGCGAGCGGCTTGGCGCGCGGATGAGAGAGAG 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGln-ValArg 259
 |||
 DB 746 GGCAGCCGGGACCCGCGACCGCTGACCCAGTGAAGAGACAGGTGGCGGAGGTGCGCG 805
 QY 259 IalysLeu---GlnGlnGlnAlaGlnGln-IleArgLeuGln-AlaGlnAlaPheGln 277
 |||
 DB 806 CCCAAGCTGAGCAGAGCCAGCCAGCCAGCATGCGCTGAGGCGCGAGGCTTCTAGGC 865
 QY 277 ArgLeuLysSerTrpPheGlnPro-----LeuValGlnAspMetGlnArgGlnTrpAl 295
 |||
 DB 866 CCGGCTCAAGAGCTGGTTCACACCCCTGGGTTGGAACGACATGAGGCGCGCCAGTGGCG 925
 QY 295 ArgLeuValGlnLysVal-GlnAlaAlaValGlyThrSerAlaIleProValProSer 315
 |||
 DB 926 CGGCTGTGGAAAAAGGTGTCAGGCTGCCGGGGCACACCGCCCTGTGTCAGC-G 984
 QY 315 spasnHis 317
 |||
 DB 985 ACAATCAT 992
 RESULT 2
 BG763371 842 bp mRNA linear EST 15-MAY-2001
 LOCUS 602735433F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
 DEFINITION mRNA sequence.
 ACCESSION BG763371
 VERSION BG763371.1 GI:14074024
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 842)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11718 row: b column: 10
 High quality sequence stop: 817.
 Location/Qualifiers
 1. 842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:4860585"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"

Db	195	GGCCAGCGCTGGGAAGCTGCGACACTGGGCTCCCTTTGGGATTACCTGGCTGGGTCAGACA	254
QY	61	LeuSerGIuGIInValGIInIuGIuLeuLeuSerSerGIInValThrGIInIuLeuArgAla	80
Db	255	CTGTCTGACAGAGGTGCAGAGAGAGCTGTCACCTCCACAGTACCAGAACCTGAGAGGGC	314
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuInIuLeu	100
Db	315	CTGATGTGACGACCACTTGAAAGAGTTGAAGGCTTACAAATTCGAACTGAGGACACTG	374
QY	101	ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuInAlaGIuAla	120
Db	375	ACCCGGGTGGCGAGAGAGACGGGGCAGCGCTGTCCAAAGAGGTGCAGGCGGGCAGGGC	434
QY	121	ArgLeuGIuAlaAspMetLysAspValArgGIuArgLeuValGIuTyrArgGIuVal	140
Db	435	CGGTGTGGGGCGGACATGGAGAGAGCTGTGGCGGCTGTGTCAGATACCGCGCGAGTG	494
QY	141	GIuAlaMetLeuGIuGIuInSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	495	CAGGCCATCTCTGGCCAGAGCACCCAGAGAGCTGCGGGTCCGCTCCACCTCGGCC	554
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180
Db	555	AAGCTGCTGAACCGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCCTGGCAGGTAC	614
QY	181	GIuAlaGIuAlaArgGIuGIuAlaGIuArgGIuLysSerAlaIleArgGIuArgLeuGIu	200
Db	615	CAGGGCGGGGGCGGAGAGGCGCCGAGCGGCGCTTCAGCCCATCCGACAGCGCCTGGGG	674
QY	201	ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGlySerLeuAlaGIuInPro	220
Db	675	CCCCTGTGTGAACAGAGGCGCGCTGGCGGCGCCACTGTGGCTCCCTGGCGGACAGCGC	734
QY	221	LeuGIuGIuArgAlaGIuAlaIleThrLysGIuArgLeuArgAlaArgMetGIu---GIuMet	239
Db	735	CTACAGAGAGCGGGCCAGGCGCTGTGGGGCGAGCGGCTGCGCGCGATTGGAGGCACATTG	794
QY	239	LysLeuArgThrArgAspArgLeuAspGIuValLysGIuGIuInValAlaGIuValArg	259
Db	795	GGGCGAGCGCGAGCCCGAGAACCGCTGTGAGAGGTGAGAGCAGGTCGCGAAGTTGGCCG	854
QY	259	LalYsLeuGIuGIuGIuInAlaGIuGIuInIleArgLeuGIuInAlaGIuInAlaPheGIuAlaArgL	279
Db	855	GCCAGAGCTGGGGGACACAGGCGCCGCAATTGCTGTGAGGCGCAA-GCCTTCCAGGCGGCT	913
QY	279	eulYsSerTrpPheGIu 284	
Db	914	CCAAAGAGCTGCTTCAAA 930	
RESULT 4			
LOCUS	Bg761746	938 bp	mRNA linear EST 15-MAY-2001
DEFINITION	602717942P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:484141L 5',		
ACCESSION	Bg761746		
VERSION	Bg761746.1	GI:14072399	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 938)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC/DC/DTF		
	cDNA Library Preparation: Lihg Hong/Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		

found through the I.M.A.G.E. Consortium/DLNL at
http://image.llnl.gov
Plate: ILICM1674 row: c column: 12
High quality sequence stop: 767.

```

source
1. .938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4841411"
/clone_1ib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

Alignment Scores:		
Pred. No.:	8.72e-117	938
Score:	1208.00	271
Percent Similarity:	92.28%	Matches: 271
Best local Similarity:	90.94%	Conservative: 4
Query Match:	76.41%	Mismatches: 12
DB:	12	Indels: 11
		Gaps: 2

QY	1	MelLysValLeuThrPAlaIaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	57	ATGAAGGTTCTGTGGGCGCTGTCTGGTCAATTCCTGGAGGATGGCAGGCCAAGGTG	110
QY	21	GlulGlnAlaValIgluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	117	GAGCAAGCGGTGGAGACAGACAGACCAGGCCAGCTGCCCCAGACAGACCGAATGGCAGAGC	170
QY	41	GlyGlnArgTrpGluLeuAlaIaLeuGlyLysArgPheTrpAspTyrLeuArgTyrPvalIgluThr	60
Db	177	GGCCAGCGCCTGGGAACATGGCACTGGGTGCTCTTGGGATTCCTGGCTGGGTGCACAGCA	230
QY	61	LeuSerGluGlnValGlnGluLeuLeuLeuSerGlnValIThrGlnGluLeuArgAla	80
Db	237	CTGTCTGACGAGGTGCAGGAGGAGCTGCTCACTGCTCCAGGTCACCCAGGAATGAGGGCG	290
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGluLeu	100
Db	297	CTGATGGACGAGACCATGTAAGAGATTGAAGGCTTACAAATCGGAATCGAGAGAACACTG	350
QY	101	ThrProValAlaGluGluThrArgAlaIaArgLeuSerLysGluLeuGlnAlaIaGlnAla	120
Db	357	ACCCGGGTGGCGAGGAGACCGCGGCAAGGCTGTCCAAAGGAGCTGCAGCGCGCCAGGCC	410
QY	121	ArgLeuGlyAlaAspMetGluAspValArgGlyLysLeuValIgluTyrArgGlyLysVal	140
Db	417	CGGCTGGGCGCGGACATGGAGAGCTGTCCGGCCGCTGGTGCAGTACCGCGGAGAGGTG	470
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	477	CAGGCCATGGCTCGGCCGACAGAGACCGAGAGAGCTGGGGGTGGCCCTGCCACTTGGCC	530
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	537	AAGTCGGTAAAGCGGCTCTCCCGGATGCCGATACCTGCAGAGAGCCCTGGCAGTGTAC	590
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	597	CAGCGCGAGGCCCGAGAGGGCGCGAGAGCGCTGTCAAGGCGCAATCCCGAGCGGCTTGGG	650

Alignment Scores:			
Pred. No.:	5,95e-115	Length:	927
Score:	1190.50	Matches:	274
Percent Similarity:	92.00%	Conservative:	2
Best Local Similarity:	91.33%	Mismatches:	14
Query Match:	75.30%	Indels:	12
DB:	12	Gaps:	1
US-09-827-854-14 (1-317) x BG472299 (1-927)			
QY	1	MettysValLeuTriPAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
DB	47	ATGAAGTTCTGTGGCTGCTGTGTCATCTCTGCGAGATGCCAAGCGCAAGTGT	106
QY	21	GlutAlaIvalaGluThrIuProGluIuPArgGlnInThrGluTrpGlnSer	40
DB	107	GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTCCGACAGACAGCGAGGAGAC	166
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
DB	167	GGCAGCGCTGGGAAGTGGCACTGGCTGCTTTGGGATTACTGGCTGGGTCAGACA	226
QY	61	LeuSerGluGlnIvalGlnGluLeuLeuSerSerGlnIvalThrGlnIuLeuArgala	80
DB	227	CTGTGTAGACAGGTGAGAGAGAGCTGCTCAGCTCCAGCTACCCAGCACTGAGAGCGC	286
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu	100
DB	287	CTGATGAGACGACCACTGAGAGAGTGTGAGAGCTTACAAATCGGAACTGGAGAGCAACTG	346
QY	101	ThrProValaIaGluGluThrArgAlaLeuSerLysGluLeuGlnAlaIaGlnAla	120
DB	347	ACCCCGGTGGGAGAGAGCGGGGACGCTGTCCAAAGAGCTGAGGCGGCGAGCGC	406
QY	121	ArgLeuGlyAlaAspMetLysPylArgGlyArgLeuValGlnTrpArgGlyGluVal	140
DB	407	CGGCTGGGCGGACATGAGAGCGTGTCCGCGCTGTGTGTCAGTACCGCGCGAGGTG	466
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
DB	467	CAGGCGATCTCTGGCCAGAGACCCGAGAGCTGGGCTGGCTGCTCCCACTGCGCGC	526
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp	180
DB	527	AAGCTGCGTAAGCGGCTCTCCGCGATGCCATGACCTCAGAAAGCCCTGGCAGGTAC	586
QY	181	Gln-AlaGlyAlaArgGluGlyAlaGluArgGly--LeuSerAlaIleArgGluArgLeu	199
DB	587	CAGGCGCGGCGCGCGAGGCGCGAGCGCGCCCTCAAGCGCATCCGCGAGCGCCTT	646
QY	200	--GlyProLeuValaGluGlnGlyArgValArgAlaIaThrValaGlySerLeuAlaGlyG	219
DB	647	GGGCGCCCTGGGTGGACAGGCGCGGCTGGGCGCGCACTGTGGCTCTCGGCGCGGC	706
QY	219	IuProLeuGlnIuArgAlaGlnAlaIaTrpGlyLysArgLeuArgAlaArgMetGluGlu	239
DB	707	ACGCG-CTACAGAGCGGCGCGAGCGCTGGGCGGAACGGTG-CGCGGCGGAGTGGAGANA	764
QY	239	etGlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnIvalaGlu-ValAr	258
DB	765	TGGGAGAGGAGAGCGCGAGCTGGACCGAGGTGAAGAGCAGGTGGCGAGAGGTGG	824
QY	258	gAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAla---PheGlnAl	277
DB	825	CCGCAAGCTGGAGAGAGCGAGCGAGCATACGGCTGCGAGGCGGAGGGGCTTCCAGGCG	884
QY	277	aArg-LeuLysSerTrpPhe-GluProLeuValaGluAspMet	290
DB	885	CCGGCTCAAGAGAGTGTTCGAGGCCCTGGGTGGAGAACATG	926
RESULT 8			
BIS97743			
LOCUS	BIS97743	922 bp	mRNA linear EST 07-SEP-2001

DEFINITION			
603248609P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',			
mRNA sequence.			
ACCESSION			
BIS97743			
VERSION			
BIS97743.1			
KEYWORDS			
GI:15490682			
SOURCE			
human.			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
1 (bases 1 to 922)			
NIH-MGC http://mhc.nci.nih.gov/ .			
AUTHORS			
National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE			
Unpublished (1999)			
JOURNAL			
Contact: Robert Strausberg, Ph.D.			
COMMENT			
Email: cgabbs-remail.nih.gov			
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki			
Toshiyuki and Piero Carninci (RIKEN)			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LNLN at:			
http://image.lnl.gov			
Plate: LLM11760 row: b column: 04			
High quality sequence stop: 782.			
FEATURES			
source			
1..922			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:5300259"			
/clone_lib="NIH_MGC_96"			
/tissue_type="hypothalamus"			
/lab_host="DHI0B"			
/note="Organ: Brain; Vector: pBluescript (modified			
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag			
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',			
size-selected for average insert size 2.3 kb and			
normalized to R0F 5. This is a primary library enriched			
for full-length clones and constructed using the			
cap-trapper method (Carninci, in preparation). Library			
constructed by M. Brownstein (NIH/NHGRI, National			
Institutes of Health). Note: this is a NIH_MGC Library."			
BASE COUNT			
176 a 273 c 359 g 114 t			
ORIGIN			
Alignment Scores:			
Pred. No.:	1.22e-114	Length:	922
Score:	1187.50	Matches:	256
Percent Similarity:	93.19%	Conservative:	4
Best Local Similarity:	91.76%	Mismatches:	15
Query Match:	75.11%	Indels:	5
DB:	13	Gaps:	1
US-09-827-854-14 (1-317) x BIS97743 (1-922)			
QY	1	MettysValLeuTriPAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
DB	73	ATGAAGTTCTGTGGCTGCTGTGTCATCTCTGCGAGATGCCAAGCGCAAGTGT	132
QY	21	GlutAlaIvalaGluThrIuProGluIuPArgGlnInThrGluTrpGlnSer	40
DB	133	GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGGCCAGACAGAGCGAGAGAC	192
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
DB	193	GGCCAGCGCTGGGAAGTGGCACTGGTGGTGGATTTACTGCTGGCTGGCGAGACA	252
QY	61	LeuSerGluGlnIvalGlnGluLeuLeuSerSerGlnIvalThrGlnIuLeuArgala	80
DB	253	CTGTGTAGACAGGTGAGAGAGAGCTGCTCAGCTCCAGAGTACCCAGAGACTGAGGCG	312
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu	100

RESULT 10
BI600563
LOCUS
DEFINITION 603244936F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
mRNA sequence.
ACCESSION BI600563
VERSION BI600563.1 GI:15493502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 811)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carpinici (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM11726 row: 9 column: 10
High quality sequence stop: 783.
location/Qualifiers
1. 811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287329"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTGTGTGTGTGTGT-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT 148 a 245 c 316 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 1,3e-113 length: 811
Score: 1177.00 Matches: 240
Percent Similarity: 97.97% Conservative: 1
Best Local Similarity: 97.56% Mismatches: 4
Query Match: 74.45% Indels: 1
DB: 13 Gaps: 0
US-09-827-854-14 (1-317) x BI600563 (1-811)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 74 ATGAGAGGTCTGTGGTGGCGCTTCTGCTGCATTCCTGGCAGATGCCAGGCCAAGGTG 133
QY 21 GUGlnAlaValGlnThrGlnProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
DB 134 GAGCAAGCGGTGAGACAGAGCCGAGCGCGCCGACACAGACGAGGAGGAGAGC 193
QY 41 GlnGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 194 GGGCAGCGGTGGAGACGTGGCTGCTTTGGGATTAACCTCGCTGGTGGACAGACA 253
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 254 CTGTCTGAGCAGGTGCAGAGGAGCTCTACAGTCCAGAGTACCCAGGAACTGAGGGCG 313

QY 81 LeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 314 CTGATGAGACGAGACCATGAGAGGAGTGAAGGCTTCAAAATCGGAAGTGGAGAACACTG 373
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 374 ACCCGGTGGCGGAGAGACGGGGCAGCGCTGTCCAGAGACTCGCAGGGCGGAGGCC 433
QY 121 ArgLeuGlnAlaAspMetGlnAspValArgGlnArgLeuValGlnTyrArgGlnLysVal 140
DB 434 CGGCTGGCGCGCGACATGAGAGACGTGTGGCGCCCTGTGTCTACATACCGCGGAGGTG 493
QY 141 GlnAlaMetLeuGlnSerThrGlnGlnGlnLeuArgValArgValAlaSerHisLeuArg 160
DB 494 CAGGCAATGTCTGGCCAGAGCACCGAGAGGTGGGGTGGCTTCCCTCCACCTGGCGC 553
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 554 AAGCTGCGTAAGCGGCTCTCCGCGATGCCATGACAGCAAGCGCTGGCACTGATAC 613
QY 181 GlnAlaGlnAlaArgGlnGlnValAlaGlnArgGlnLysSerAlaIleArgGlnArgLeu 200
DB 614 CAGGCGGGGGCCCGCAGAGGCGCCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 673
QY 200 YProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnLys 220
DB 674 CCCCTGTGGTGAACAGGGCGCGCTCGGGCCGCCACTGTGGGCTCTCTGGCAGGACGC 733
QY 220 OLeuGlnGlnArgAlaGlnAlaTrpGlnGlnArgLeuArgAlaArgMetGlnGlnMet 240
DB 734 GCTACAGAGCGGGCGCCAGCGCTGGGGCGAGCGCTGGCGGGCGGAGTGGCAATGGG 793
QY 240 YSerArgThrArgAsp 245
DB 794 CAGCCGGAACCGCGAA 809
RESULT 11
LOCUS BG829472 845 bp mRNA linear EST 22-MAY-2001
DEFINITION 602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 845)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LUCM1790 row: 9 column: 17
High quality sequence stop: 829.
location/Qualifiers
1. 845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4899112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOT87; Site.1: XhoI;

QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 253 CTGTCTGACAGAGTGAGAGAGAGCTCTCCAGCTCCAGAGTACACCCAGAGAACTGAGGGCG 312
QY 81 LeuMetaspGIuThrMetIySgluLeuYsAlaTYrIySserGIuLeuGIuGIuLeu 100
|||||
Db 313 CTGATGAGACAGACATGATGAGAGCTTGAAGGCTTCAATTCGAACTGGAGAGAACACTG 372
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerIySgluLeuGIuAlaGIuAla 120
|||||
Db 373 ACCCGGCTGGCGAGAGACCGCGGACGCTGTCCAGAGAGCTGAGGCGGCGAGGCGC 432
QY 121 ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTYrArgGIuGIuVal 140
|||||
Db 433 CGCGTGGGGCGGACATGAGAGACGTGGCGGCCCTGTGTCTACATCCGCGGCGAGGTG 492
QY 141 GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCAATGCTCGCGCAGAGCACCGAGGAGCTGCGGCTGCGCTCCACCTGGCGC 552
QY 161 LysLeuArgIySArgLeuLeuArgAspAlaAspAspLeuGIuIySArgLeuAlaValTYr 180
|||||
Db 553 AAGCTGCGTAAGCGGCTCTCCGCAATGCCATGACCTGCAGAGAGCGCTGGCAGCTGAC 612
QY 181 GIuAlaGIuAlaArgGIuGIuValArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIy 200
|||||
Db 613 CAGCGCGGGCGCGGAGAGCGGCCAGCGCGGTCTCAGCGCATCCGCGAGCGCTGGGG 672
QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIySerLeuAlaGIuInPro 220
|||||
Db 673 CCCCTGGTGGAGACAGAGCGCGCTGGCGCGCCACTGTGGCTCTCCCTGCGCAGCGCGG 732
QY 221 LeuGIuGIuArgAlaGIuAlaThrPGLIyGIuArgLeuArgAlaArgMetGIuGIuMet 239
|||||
Db 733 CTACAGAGAGCGGGCCAGAGCTGGGGCGAGCGGTG-CGCGCGGATGAGAGAGATG 788
RESULT 13
B1551475 919 bp mRNA linear EST 05-SEP-2001
LOCUS 603194314F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION B1551475
VERSION B1551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L14M1691 row: 1 column: 04
High quality sequence stop: 812.
Location/Qualifiers
1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptPR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 2.5e-109 Length: 919
Score: 1137.00 Matches: 246
Percent Similarity: 93.26% Conservative: 3
Best Local Similarity: 92.13% Mismatches: 17
Query Match: 71.92% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-14 (1-317) x B1551475 (1-919)
QY 1 MetIySValLeuThrPGLIAlaLeuLeuValThrPheLeuAlaGIyCySgluAlaIySVal 20
|||||
Db 73 ATGAAGTTCGTGTGGCTGCGCTTCTGCTACATTCCTGGCAGATGCCAGGCAAGGTG 132
QY 21 GIuGIuAlaValGIuThrGIuProGIuProGIuLeuArgGIuGIuThrGIuThrPGLIInser 40
|||||
Db 133 GAGCAAGCGGTGAGACAGAGCCGGAGCCGAGCTGCGCCACAGACAGAGAGGAGAGC 192
QY 41 GIyGIuArgTTPGLIuLeuAlaLeuGIyArgPheTTPAsPTyLeuArgTTPValGIuThr 60
|||||
Db 193 GGCCAGCGCTGGAGACTGGCACTGGGTGCTTTTGGAFTTACTCGCTGGGTGAGACA 252
QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 253 CTGTCTGACAGAGTGACAGAGAGAGCTGCTCCAGAGTACCCAGAGAACTGAGGGCG 312
QY 81 LeuMetaspGIuThrMetIySgluLeuYsAlaTYrIySserGIuLeuGIuGIuLeu 100
|||||
Db 313 CTGATGAGACAGACATGATGAGAGCTTGAAGGCTTCAATTCGAACTGGAGAGAACTG 372
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerIySgluLeuGIuAlaGIuAla 120
|||||
Db 373 ACCCGGCTGGCGAGAGACCGCGGACGCTGTCCAGAGAGCTGAGGCGGCGAGGCGC 432
QY 121 ArgLeuGIuAlaAspMetGIuAspValArgGIuArgLeuValGIuTYrArgGIuGIuVal 140
|||||
Db 433 CGCGTGGGGCGGACATGAGAGACGTGGCGGCCCTGTGTCTACATCCGCGGCGAGGTG 492
QY 141 GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCAATGCTCGCGCAGAGCACCGAGGAGCTGCGGCTGCGCTCCACCTGGCGC 552
QY 161 LysLeuArgIySArgLeuLeuArgAspAlaAspAspLeuGIuIySArgLeuAlaValTYr 180
|||||
Db 553 AAGCTGCGTAAGCGGCTCTCCGCAATGCCATGACCTGCAGAGAGCGCTGGCAGCTGAC 612
QY 181 GIuAlaGIuAlaArgGIuGIuValArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIy 200
|||||
Db 613 CAGCGCGGGCGGACATGAGAGACGTGGCGGCCCTGTGTCTACATCCGCGGCGAGGTG 672
QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIySerLeuAlaGIuInPro 220
|||||
Db 673 CCCCTGGTGGAGACAGAGCGCGCTGGCGCGCCACTGTGGCTCTCCCTGCGGCGAGCTT 732
QY 221 LeuGIuGIuArgAlaGIuAlaThrPGLIyGIuArgLeuArgAlaArgMetGIuGIuMetGIy 240
|||||
Db 733 A--CAGAGAGCGGGCGAGAGCTGGGGCGAGCGGTGCGGCCCGGATGAGAGAAATGGG 790
QY 240 YSerArgThrArgAspArgLeuAspGIuValIySgluGIuValAlaGIuValArgAlaIy 260
|||||
Db 791 CAGCGGAGCCGGGACG--CTGGAGAGAGTGAAGAGCAGTGGGGGAGGG--TGGGCCAA 847

QY 260 sleugluglunlagln 266
 |||||:::||||
 Db 848 GCTGAGACACACAGAG 866

RESULT 14
 LOCUS BM728696 706 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-EJ0-aiu-1-08-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-aiu-1-08-0-UI 5', mRNA sequence.
 ACCESSION BM728696
 KEYWORDS BM728696.1 GI:19050022
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse
 FEATURES
 source Location/Qualifiers
 1..706
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EJ0-aiu-1-08-0-UI"
 /clone_1lb="UI-E-EJ0"
 /tissue_type="fetal eyes, lens, eye anterior segment,
 optic nerve, retina, Retina Foveal and Macular, RPE and
 Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EJ0 is a subtracted cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are: fetal eyes, AGATCAGAG
 ; lens, CGATTACGCA; eye anterior segment, AATGCCGCAT;
 optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
 Macular, GTCC; RPE and Choroid, ACCCA. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."

BASE COUNT 123 a 218 c 283 g 82 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.2e-109 Length: 706
 Score: 1136.00 Matches: 233
 Percent Similarity: 98.73% Conservative: 0
 Best Local Similarity: 98.73% Mismatches: 1

Query Match: 71.85% Indels: 2
 DB: 14 Gaps: 0
 US-09-827-854-14 (1-317) x BM728696 (1-706)

QY 15 GlycylGlnAlaLysValGluGlnAlaVal-GluThrGluProGluLeuArgG1 34
 |||||
 Db 2 GGATGCCAGGCCAAGGTGGAGCAAGCGGTGAGAGACCGGACCCGAGCTGGCCA 61
 QY 34 ngIntnrgluTrpGlnSerGlyGlnArgTrpGluAlaLeuGlyArgPheTrpAspTy 54
 |||||
 Db 62 GCAGACCGAGTGGCAGAGCGCGCAGCGCTGGAGACTGGCAGCTGGCTTTGGATT 121
 QY 54 rleuArg-TrpValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnV 74
 |||||
 Db 122 CTTGCCCGTGGGTCAGACACTGTGTGACAGAGTGCAGAGAGAGCTGTCTACCTCC 181
 QY 74 alThrGlnGluLeuAlaGlnAlaLeuMetAspGluThrMetLysGluLeuLysAlaTy 94
 |||||
 Db 182 TCACCCAGGAACAGAGGCGGTGATGACGAGACCATGGAAGGATTGAAGCCCTACAC 241
 QY 94 erGluLeuGluGlnGlnLeuThrProValAlaGluGlnThrArgAlaArgLeuSerLys 114
 |||||
 Db 242 CGGAACCTGGAGAACAACTGACCCCGGTGGCGAGAGACCGCGGCTGTCCAAAG 301
 QY 114 lueGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAspValArgGlyArgLeu 134
 |||||
 Db 302 AGCTGAGCGCGCGCAGCGCCGCTGGCGGACATGAGAGACGTGCGCGCCGCTGG 361
 QY 134 alGlnTrpArgGlyGluValGlnAlaMetLeuGlnSerThrGlnGluLeuArgVal 154
 |||||
 Db 362 TGGAGTACCGCGCGGAGGTGAGGCGCATGTGCGCCAGACACCGGAGAGCTGGGTTCC 421
 QY 154 rgluAlaSerHisLeuArgLysLeuArgLysArgLeuArgAspAlaAspAspLeuG 174
 |||||
 Db 422 GCGTGGCTCCACACCTGGCCGCAAGCTCGTAAGGGCTCTCCGCGCATGGCATGAC 481
 QY 174 lnlYsArgLeuAlaValTrpGlnAlaGlyAlaArgGluGluAlaArgGlyLeuSer 194
 |||||
 Db 482 AGAAGCGCTGGCAGTGTACAGAGCGCGGCGCCGACAGGCGCGCGCTCTCAGCG 541
 QY 194 lalLeuArgGluArgLeuGluProLeuValGluGlnGlnArgValArgAlaAlaThrVal 214
 |||||
 Db 542 CCATTCGCGAGCGCTGGGGCCCCCTGTGTGAAACAGGGCCGCGCGGCGCACTGTGG 601
 QY 214 lYserLeuAlaGlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArg 234
 |||||
 Db 602 GCGTCTGGCGCGCGCGCGCTACAGAGCGGGCCAGCGCTGGCGAGCGGCTGGCGG 661
 QY 234 laArgMetGluGluMetGlySerArgThrArgAspArgLeuAsp 248
 |||||
 Db 662 CCGGATGAGAGATGGGCGACCGCGAGCCGCGAGCGCTGGAGC 705

RESULT 15
 LOCUS AK010261 1110 bp mRNA linear HTC 19-JAN-2002
 DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
 clone:2400003C12.apolipoprotein E, full insert sequence.
 ACCESSION AK010261
 KEYWORDS AK010261.1 GI:12845571
 SOURCE HTC, CAP trapper.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636


```
Db 523 AACATGCCCAAGCCCTTGATGCGGGATGCCAGATCTGCCAAGCCCTAGCTGTCTAC 582
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 583 AAGCAGGGGAGCGCTAGGGGCGCGAGCGCGGTGTGAGTGCCATCCGTGAGCGCCTGGGG 642
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
   ||::::::::::::::::::::::::::::|||:::||||::||| |||||
Db 643 CCTCTGGTGGAGCAAGGTGCGCACGCTGCCAACCTAGGCGCTAGGCGCGCCAGAGCT 702
QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
   ::::: ||::::::::::::::::::::|||:::||||| ||::::::::::::|||
Db 703 ATGTGCGATGCGCGCCAGCGCTTTGGTGACCGCATCCGAGGGCGGCTGGAGCAAGTGGGC 762
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
   :::::: ||::::::::::::::::::::|||::: |||::: ||::::|||
Db 763 AACCAAGCGCCGTGACCGCTAGAGAGAGTGGGTGACACATGGAGCGAGTGGCGCTCCAG 822
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
   ::::::::::||| ||::::::::::::::::||| ||| ||| ||| ||| |||
Db 823 ATGGAGGAGACAGACCCCAATATACGCTGACGCGGAGATCTTCCAGGCCCGCTCAAG 882
QY 281 SerTrpPheGluProLeuValGluAspMetGluArgGlnTrpAlaGlyLeuValGluLys 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 883 GGCTGTTCCAGCCCAATAGTGAGACATGCATGCCAGTGGGCAAACTGATGGAAGAAG 942
QY 301 ValGlnAlaAlaValGlyThrSer-----AlaAlaProValProSerAspAsn 316
   ::::::::::||| |||::: ||| ||| ||| ||| ||| ||| ||| |||
Db 943 ATACAGGCGCTGTGTGCTACCAACCCCATCATCACCCAGTGGGCCAGAGGAAT 996
```

Search completed: March 14, 2003, 20:13:47
Job time : 1296.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALVTLFLGCOAKV.....VEKVQAVGTSAPVPSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/rnact_11032003_101610_27486/app_query.fasta.1.3576
-DB=GenEmbl -QFMT=fastlap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.cgn.1.13008.rnact_11032003_101610_27486 -NCPU=6 -ICPU=3
-MARK=NO_MMAP -LARGQUERY=NEC_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-YGAPOP=10 -YGAPEXT=0.5 -DELPEXT=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1585	100.0	1110	6 E00359	E00359 cDNA coding
2	1585	100.0	1110	6 E00823	E00823 DNA sequence
3	1585	100.0	1147	6 AX302545	AX302545 Sequence
4	1585	100.0	1156	6 BD004278	BD004278 Apo E hum
5	1585	100.0	1156	9 HUMAPOE3	K00396 Homo sapien
6	1585	100.0	1186	9 BC003557	BC003557 Homo sapi
7	1578	99.6	1110	6 E08423	E08423 DNA coding
8	1577	99.5	1156	6 BD004277	BD004277 Apo E hum
9	1576	99.4	1157	6 AX409597	AX409597 Sequence
10	1576	99.4	1157	6 I15975	I15975 Sequence 1
11	1576	99.4	1157	6 HUMAPOE	M12529 Human apoli
12	1576	99.4	1157	6 HUMAPOE	BD004279 Apo E hum
13	1573	99.2	1156	6 BD004279	X13887 Monkey mRNA
14	1451	91.5	1178	9 MFAPOE	AF261279 Homo sapi
15	1396.5	88.1	5491	6 AF261279	AF261279 Sequence
16	1396.5	88.1	41907	6 AX358722	AP050154 Homo sapi
17	1396.5	88.1	41907	9 AF050154	AC011481 Homo sapi
18	1396.5	88.1	107567	9 AC011481	M10065 Human apoli
19	1379	87.0	5515	9 HUMAPOE4	M23322 Baboon apoli
20	1350	85.2	5413	9 AF261280	AF261280 Pan trogl
21	1267	79.9	4762	9 BABAPOE	M23322 Baboon apoli
22	1266	79.9	208239	2 AC021988	AC021988 Homo sapi
23	1258	79.4	1138	4 AF303830	AF303830 Tupaia gl
24	1170	73.8	1060	4 RABAPOE	M36603 Rabbit apoli
25	1163.5	73.4	965	6 AX384545	AX384545 Sequence
26	1163.5	73.4	1108	4 BTPAPOE	X61171 B. taurus MR
27	1163.5	73.4	5617	6 AX384541	AX384541 Sequence
28	1163.5	73.4	6026	6 AX384539	AX384539 Sequence
29	1151	72.6	718	9 AF20049753	AF200499 Pan trogl
30	1149	72.5	1045	10 MUSAPOE	M12414 Mouse apoli
31	1149	72.5	1104	10 BC028816	BC028816 Mus muscu
32	1147	72.4	718	9 AF20050053	AF200502 Gorilla g
33	1147	72.4	718	9 AF20050053	AF200505 Pongo pyg
34	1145.5	72.3	1154	4 BTPAPOE	X64839 B. taurus MR
35	1143	72.1	1122	4 AF20050653	AF200508 Hylobates
36	1143	72.1	1122	4 SSAPOE	AF200508 Hylobates
37	1138	71.8	1126	6 AR164342	AR164342 Sequence
38	1138	71.8	1126	6 AR205885	AR205885 Sequence
39	1094	69.0	951	10 S76779	S76779 rAPOE-apol1
40	1072	67.6	959	10 MUSAPOE	M73480 Mus muscu
41	1056.5	66.7	1069	10 RATPOE	J00705 rat apolipo
42	1029	64.9	228698	2 AC127479	AC127479 Mus muscu
43	1029	64.9	237653	2 AC073760	AC073760 Mus muscu
44	1003	63.3	4856	10 MUSAPE	D00466 Mus muscu
45	1002	63.2	4267	6 AR164387	AR164387 Sequence

RESULT 1

ALIGNMENTS

LOCUS	E00359	1110 bp	RNA	linear	PAT 29-SEP-1997
DEFINITION	cDNA coding human apolipoprotein E3.				
ACCESSION	E00359				
VERSION	E00359.1	GI:2168646			
KEYWORDS	JP 1985118189-A/1.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Primates; Carnivora; Homnidae; Homo. (bases 1 to 1110)				
AUTHORS	Terasanishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.				
TITLE	DNA FRAGMENT				
JOURNAL	Patent: JP 1985118189-A 1 25-JUN-1985;				

COMMENT

OS	human
PN	JP 1985118189-A/1
PD	25-JUN-1985
FP	29-NOV-1983 JP 1983224980
PI	TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
PI	KIMURA MASAKO,
PC	IKEDA YASUKO
CC	C12N15/00,C07H21/04//C12P21/00;
CC	strandedness: Double;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
CC	*source: tissue_type=lyver;
FH	location/Qualifiers
FH	
FT	CDS
FT	15..968
FT	/product='apolipoprotein E3 precursor' FT
sig_peptide	15..68
FT	/product='apolipoprotein E3 signal peptide' FT
mat_peptide	69..965
FT	/product='apolipoprotein E3'.

```

FEATURES
source      Location/Qualifiers
1. .1110

```

BASE COUNT	198 a	353 c	416 g	143 t
ORIGIN				

Alignment Scores:	
Pred. No.:	9.31e-81
Score:	1585.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	1110
Matches:	317
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-827-854-15 (1-317) x E00359 (1-1110)

QY	1	MeysvVAlleuTPpAlAlAlleuValIthrPheleuAlaGICysGlnAlaIstAl	20
Db	15	ATGAGGCTTCTGTGGCTGCCCTTGCTGCATTTCTGCCAGGATGCCAGCAAGTG	74
QY	21	GlucInAlaValAlGluThrGluProGluProGluLeuAaGrgInIthrGluTrpInser	40
Db	75	GAGCAAGCGGCTGGAGACAGACCCGAGCCAGCTGCCAGCAGACCGACTGCCAGAGC	134
QY	41	GlycInAlaTrpGluLeuAlaLeuGluValAaPheTrpAspTyrLeuAaGtrValGluThr	60
Db	135	GGCCAGCGCTGGAACTGCACTGGGTGCTTTGGGGATTACCTGGCGTGGTCCAGCA	194
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValIthrGlnGluLeuAaAla	80
Db	195	CTGTCTAGACGAGTGCAGAGAGACTGCTGCTACCTCCAGTCACCCAGAACCTAGGGCGC	254
QY	81	LeuMetAspGluTrpMetLeuLysGluLeuLysAlaTyrIysSerGluLeuGlnGlnLeu	100
Db	255	CTGATGACGACCACTGAGAGAGATTGAAGCCCTPACAAATGTGAACGTGAGAACCACTG	314

QY	101	THRPVQVALIAGLUGLUGLHRAIRGALAAARGLEUSERLYGILUENCLINALAAGLALNA	120
Db	315	ACCCCGGTGGCGAGAGACGCCGGCAGCCGTCTGTCCAAAGAGCTGCAGCGCGCCAGGCC	374
QY	121	ACGLEUGLYALAApmetGLuAspVALyCSylARgLEuVALGLNtyrARGLyGLuVAL	140
Db	375	CGGGTGGCGCGGACATGAGAGACGTGTGCGGCCGCTGTGCAGTACCGCGGCGAGGTG	434
QY	141	GLNALmetLEuGLyGLNserThrGLUGLULeuARgVALARgLEuALAserHisLEuARg	160
Db	435	CAGGCATCTCTGGCGCAGACACCGAGGAGCTGCGGGGTGCGCTGCCCTCCACCTCGGCC	494
QY	161	LYSLeuARgLYARgLEuLEuARgASpVALAspAspLEuGLNlyARgLEuVALtyr	180
Db	495	AAGCTGCGTAAACGGCTCTCCGCGATGCGCGTGAACCTGCACAAAGCGCTTGCAGGTAC	554
QY	181	GLNALAGLYALAArGLUGLYALAGLUAARGLYLEuSERALILEARGLUARgLEuGLY	200
Db	555	CAGGCCGGGGGGCCCGGAGGGGCCCCAGCGCGGCTTCAGGCCATCCGACGCGCTGGGG	614
QY	201	ProLEuVALGLUGLUGLYARgVALARgALAArHVALyGLYSerLEuALAGLYGLNpro	220
Db	615	CCCTGTGTGGAACAGAGGCCGCGTGGGGGCGCCACATGTGGGCTCCCTGGCGGCGAGCGG	674
QY	221	LEuGLUGLYARgALIGLINALtrpGLyGLUARgLEuARgALAArGmetGLUmetGLY	240
Db	675	CTACAGAGAGCGGGCCAGGCCCTGGGGCGAGCGCGTGGCGCGCATGAGAGATGGCG	734
QY	241	SerATgTYrARgASpARgLEuAspGLUVALySGLUGLINALAGLYALAArGLALyS	260
Db	735	AGCCGGACCCCGACCGCTGTGACACAGGTGAGAGAGACAGTGGCGAGGTGGCGCCAGG	794
QY	261	LEuGLUGLUGLINALAGLGNILleARgLEuGLINALAGLUALAphGLINALARgLEuLYS	280
Db	795	CTGGAGGAGAGAGCCAGACAGATACGCTGCAGGCGGAGGCTTCAGAGCCCGCTCAAG	854
QY	281	SerTPHleGLUProLEuVALGLUAspmetGLNARgGLINTPAlAGLYLEuVALGLyS	300
Db	855	AGCTGTTCGAGACCCCTGTGTGAAACATATGCAGCCCACTGGTGGCGCGGTGTGTGAGAG	914
QY	301	VALGlnALAAVALAGLYThrSerLALAAProVALProSerASPAspHis	317
Db	915	GTGACGGCTGCCGTGGCACCGCCCGCCCTGTGCCCCAGCGACAAATCAC	965
RESULT 2			
E00823		1110 bp	DNA linear
LOCUS			PAT 29-SEP-1997
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal peptide.		
ACCESSION	E00823		
VERSION	E00823.1	GI:2169084	
KEYWORDS	JP 1986096997-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1110)		
AUTHORS	Teranishi,Y., Matsui,Y., Ikeda,Y. and Kimura,M.		
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN		
JOURNAL	Patent: JP 1986096997-A 1 15-MAY-1986;		
COMMENT	MITSUBISHI CHEM IND LTD		
OS	Human (Homo sapiens)		
PN	JP 1986096997-A/1		
PD	15-MAY-1986		
PF	16-OCT-1984 JP 1984216987		
PI	TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO		
PC	C12P21/00 A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12N1:19),		
PC	(C12N15/00,		
PC	C12N1:19);		
CC	strandedness: Double;		
CC	topology: Linear;		
CC	hypothetical: No;		
CC	anti-sense: No;		
CC	*source: tissue_type=Liver;		

Qy	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	735	AGCCGGAGCCCGACGCCCTGGAGACGAGTGAAGAGCAGGTGGCGAGGAGCGCCGCAAG	794
Qy	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnIleArgLeuLys	280
Db	795	CTGAGAGAGAGCGCCAGCAGATTCGCTCCAGCGAGCGCTTCCAGGCGCGCTCAAG	854
Qy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300
Db	855	AGCTGCTTCAGAGCCCTGTGTGGAAAGATACAGCCAGTGGCGCGCTGTGTGGAGAAG	914
Qy	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	915	GTGCAGGCTGCGGTGGGCACACGCGCGCGCTGTGCTCCAGCAGCAATCTAC	965
RESULT 3			
LOCUS	AX302545	1147 bp	DNA
DEFINITION	Sequence 63 from Patent WO0175177.		
ACCESSION	AX302545		
VERSION	AX302545.1	GI:17383082	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Morin, P. J., Sherman-Baust, C. A., Pizer, E. S. and Hough, C. D.		
TITLE	Tumor markers in ovarian cancer		
JOURNAL	Patent: WO 01/75177-A 63 11-OCT-2001;		
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)			
FEATURES			
source	1. 1147	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	210 a	365 c	425 g 147 t
ORIGIN			
Alignment Scores:			
Pred. No.:	9.64e-81	Length:	1147
Score:	1585.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-15 (1-317) x AX302545 (1-1147)			
Qy	1	MetyValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	46	ATGAAGTTCTGTGGCTGCTGCTGTCACATTCTGCGAGAGATGCCAGGCGCAAGTG	105
Qy	21	GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnIleThrGluTrpGlnSer	40
Db	106	GAGCAAGCGGTGGAGACAGACGCGAGCCGAGACTGCGCAGCAGACCCAGTGGCAGACG	165
Qy	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	166	GGCCAGCGCTGGAACTGCGACATGGGCTGCTTTGGGATTAACCTGGCTGGGGTGCAGACA	225
Qy	61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	226	CTGTCTGAGCAGGTGCAGAGAGAGCTCTAGCTCCAGGTACCCAGAACTGAGAGCGG	285
Qy	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu	100
Db	286	CTGATGAGCAGACCATTAAGAGAGATTGAAGCTTCAAAATCGGAATCGAGAAACAATG	345
Qy	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnIleAlaGlnAla	120
Db	346	ACCCCGGTGGCGAGAGCGCGGCGAGCTGTCCAAAGAGAGCTGAGCGCGCGAGGCC	405

Qy	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaVal	260
Db	735	AGCCGACCCCGGACCGCTGGACAGGTGAAGAGCGAGTGGCGGAGTGGCGCCCAAG	794
Qy	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
Db	795	CTGGAGAGACAGGCCACACAGATACGCTGCAGGGCGGAGGCTTCAGAGCCGCTCAAG	854
Qy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpPalaGlyLeuValGluLys	300
Db	855	AGCTGTGTGAGCCCTGCTGGAGACATGACAGCCCGAGTGGCGCGGCTGTGGAGAG	914
Qy	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	915	GTGAGGCTCGGTGGGCAACGCGCCCGCTGTGCTCCAGCAGCAATCAC	965
RESULT 3			
LOCUS	AX302545	1147 bp	DNA linear PAT 30-NOV-2001
DEFINITION	Sequence 63 from Patent WO0175177.		
ACCESSION	AX302545		
VERSION	AX302545.1	GI:17383082	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Morin, P.J., Sherman-Baust, C.A., Pizer, E.S. and Hough, C.D.		
TITLE	Tumor markers in ovarian cancer		
JOURNAL	Patent: WO 0175177-A 63 11-OCT-2001;		
FEATURES	THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)		
Source	1..1147		
Location/Qualifiers	/organism="Homo sapiens"		
BASE COUNT	210 a 365 c 425 g 147 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	9.64e-81	Length:	1147
Score:	1585.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-15 (1-317) x AX302545 (1-1147)			
Qy	1	MetyValLeuThrPalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db	46	ATGAAGGTTCTGTGGCTGCTGCTGTCACATTCTGCGAGATGCGCAGGCTG	105
Qy	21	GlulAlaValAlaGluThrGluProGluProGluLeuArgGlnThrGluTrpIleu	40
Db	106	GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCGACAGACAGCGATGGAGAGC	165
Qy	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	166	GGCCAGCCCTGGAACTGCTGCTGCTTTGGATTAACCTGCTGGGTGCAGACA	225
Qy	61	LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla	80
Db	226	CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGAACTGAGGGCG	285
Qy	81	LeuMetAspGluThrMetIleGluLeuLysAlaTyrIleSerGluLeuGluGlnLeu	100
Db	286	CTGATGAGCAGACCATTAAGAGAGTGAAGCCCTACAAATCGAATCGAGAGCAACTG	345
Qy	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla	120
Db	346	ACCCCGGTGGCGAGAGACCGCGGCGAGCTGTCCAAAGAGCTGCAAGCGCGAGGCC	405

QY 121 ArgLeuGIyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 406 CGGCTGGGGCGGACATGAGACGTGTGCGCGCTGTGTGATACCGCGGAGAGTTC 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 466 CAGGCGCATCTCTGGCCAGACACCGAGAGCTCGGGTGGCTCTCCCTCCACCTGCGGC 525
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 526 AAGCTCTGGTAAGGGCTCTCTCGCGATGCGGATGACCTTCAGAAAGCGCTCGGAGATGAC 585
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||
Db 586 CAGCGCGGGGGCGCGAGGGGGCGCGAGCGCGCTGACAGCGCATCCGCGAGCGCTCGGGG 645
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
|||||
Db 646 CCCCTGGTGGAAACAGGGCGCGCGCGCGCGCGCTGCTGGCGCTCCCGCGCGCGCGCG 705
QY 221 LeuGlnGlnArgAlaGlnAlaThrProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
|||||
Db 706 CTACAGAGAGCGGGCGCGAGCGCTGGGGCGAGCGGCTGCGCGCGGATGGAGATGGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
|||||
Db 766 ACGCGGACCGCGAGCGCGCTGGGGCGAGCGGCTGCGCGCGGATGGAGATGGGC 825
QY 261 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
|||||
Db 826 CTGGAGAGACAGCGCGCGAGATACCGCTGGCGAGCGCGCTTCACAGCGCGCGCGCGCAAG 885
QY 281 SerThrProGlnProLeuValGlnAspMetGlnArgGlnTyrAlaGlnGlnGlnGlnGln 300
|||||
Db 886 AGCTGCTTGGAGCCCTGTGTGAGACATGACAGCGCGCATGTGGCGCGCTGTGGGAGAG 945
QY 301 ValGlnAlaAlaValGlnGlnSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 946 GTGCAGAGCTGCGGTGGGACACAGCGCGCGCTGTGCGCGAGCACAATAC 996
RESULT 4
BD004278 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004278 Apo E humanized mammal.
DEFINITION BD004278
ACCESSION BD004278.1 GI:18632239
VERSION BD004278.1 GI:18632239
KEYWORDS JP 2001017028-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
Apo E humanized mammal
Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09, C07K14/775,
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC
FEATURES
Source FT Key CDS Location/Qualifiers
1..1156 Location/Qualifiers
/db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
ORIGIN
Alignment Scores: 9.71e-81 Length: 1156
Pred. No.: 317
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-15 (1-317) x BD004278 (1-1156)
QY 1 MetLysValLeuTyrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 61 ATGAAGATTCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 21 GlnGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
|||||
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGAGCTGCGCGAGCGCGAGCGCGAGCGAGCC 180
QY 41 GlyGlnArgThrGlnLeuAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
|||||
Db 181 GGCGACCGCTGGGAACTGGCACTGGGCTGCTTTGGGATTTACTGCGCTGGTGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
|||||
Db 241 CTGTGTGACAGAGTGCAGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnGlnGlnGlnGlnGlnGln 100
|||||
Db 301 CTGATGGAGAGACCAATGAAAGAGGTGAAGGCGCTTAACAAATCGGAATCGAGAGAAACATG 360
QY 101 ThrProValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
|||||
Db 361 ACCCGCGTGGCGAGAGACGCGGGGACGCGCTGTCCAAAGACTGCGAGCGCGCGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGGCTGGGGCGGACATGAGAGACGATGTGCGCGCGCTGTGACAGTACCGCGCGCGAGGTC 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
|||||
Db 481 CAGGCGCATCTCTGGCCAGACACCGAGAGCTGCGGGTGGCTGCGCTGCCCTCCACCTGCGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 541 AAGCTGCGTAAGCGGCTCTCCCGCATGCGCATGACCTGCAGAGCGCTGCGCATGATGCGC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||
Db 601 CAGGCGGGGGCGCGAGGGCGCGCGCGCGCGCGCTTCACGCGCATCGCGCGAGCGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
|||||
Db 661 CCCCTGGTGGAAACAGGGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
|||||
Db 721 CTACAGAGAGCGGGCGCGAGCGCTGGGGCGAGCGCGCTGCGCGCGCGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
|||||
Db 781 ACGCGGACCGCGAGCGCGCTGAGACGAGTGAAGAGACAGTGGCGAGGTGGCGCGCGCAAG 840
QY 261 LeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
|||||
Db 841 CTGGAGAGACAGCGCGCGAGATACCGCTGCGAGCGCGAGCGCTTCACAGCGCGCGCTCAAG 900
QY 281 SerThrProGlnProLeuValGlnAspMetGlnArgGlnTyrAlaGlnGlnGlnGlnGln 300
|||||
Db 901 AGCTGCTTGGAGCCCTGTGTGAGACATGACAGCGCGCGCTGTGTGAGAGAG 960

Oy	301	VglnaaiaalaValaIGlythrserlialaprovlaProsearaSPASHis 317
Db	961	GTCAGGCTGCCGTGGCACCACGCCCCCGCTGTGGCCAGCACATCAC 1011
RESULT 5	HUMAPOE3	1156 bp mRNA linear PRI 24-NOV-2000
LOCUS	HUMAPOE3	
DEFINITION	Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.	
ACCESSION	K00396	
VERSION	K00396.1 GI:178850	
KEYWORDS	apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 355 to 1156) Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W., Lofquist-Kahl,F., Karathanasis,S.K. and Zannis,V.I.	
TITLE	Identification and DNA sequence of a human apolipoprotein E cDNA clone	
JOURNAL	J. Biol. Chem. 257 (24), 14639-14641 (1982)	
MEDLINE	83082756	
PUBMED	6897404	
REFERENCE	2 (bases 250 to 777) Wallis,S.C., Rognes,S., Gill,L., Markham,A., Edge,M., Woods,D., Williamson,R. and Humphries,S.	
AUTHORS	The isolation of cDNA clones for human apolipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues	
TITLE	EMBO J. 2 (12), 2369-2373 (1983)	
JOURNAL	84131952	
MEDLINE	6199196	
PUBMED	3 (bases 1 to 1156) Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.	
REFERENCE	Synthesis, intracellular processing, and signal peptide of human apolipoprotein E	
AUTHORS	J. Biol. Chem. 259 (9), 5495-5499 (1984)	
TITLE	4 (bases 88 to 1156) McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,T.M.	
JOURNAL	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant	
MEDLINE	J. Biol. Chem. 259 (10), 6498-6504 (1984)	
PUBMED	84212473	
REFERENCE	5 (bases 577 to 624) Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.	
AUTHORS	Isolation and characterisation of a variant allele of the gene for human apolipoprotein E	
TITLE	Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)	
JOURNAL	85279526	
MEDLINE	2992507	
PUBMED	6 (sites) Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.	
REFERENCE	Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant	
AUTHORS	J. Clin. Invest. 83 (4), 1095-1101 (1989)	
TITLE	[1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4]	
JOURNAL	89198059	
MEDLINE	2539388	
PUBMED		
COMMENT		

```

epsilon-3 and epsilon-2. This sequence appears to be of the
epsilon-3 allele. [1] argues that the apo E polymorphism involves
mutations in the structural coding region; for example the
epsilon-2 phenotype which is characterized by hyperlipoproteinemia
is thought to result from a c to t change (arg to cys) at base 586
below [3]/[5]. The sequence shown is 578 homologous with human apo
A-I and 818 homologous with rat apo E. For the epsilon-4 sequence,
see the separate entry.
[U. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised
relative to [1] to record the
revisions in the FEATURES table. The sequence below is that of [U.
Biol. Chem. 258, 11422-11422 (1983)]
and [3].
Apo E is located on chromosome 19 --Jackson, Bruns and Breslow ,
PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to
the Apo C-II gene (see separate entry).
The two mutations causing type III hyperlipoproteinemia
(apolipoprotein E phenotype E3/3) produces substitutions of Arg for
Cys at amino acid 112 and Cys for Arg at amino acid position 142.
Drat entry and printed copy of sequence for [1] kindly provided by
L.L. Gyll, University of Aberdeen.
Complete source information:
Human liver [1], [U. Biol. Chem. 258, 11422-11422 (1983)], [2], [3],
[4] and blood [5], cDNA to mRNA.
Location/Qualifiers
1. .1156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.2"
/tissue_type="liver and blood"
1. .1156
/gene="APOE"
61. .1014
/gene="APOE"
/codon_start=1
/product="preapolipoprotein E"
/protein_id="AAB5546.1"
/db_xref="GI:178851"
/db_xref="GDB:600-119-691"
/translatability="MKVLMAALLVTFILAGCOAKVEAVETEPEPELROOFEMSGGRPM
ELALRFMDYLRBMVOTLSEVOEELISSQVONELRALMDERMKELKAYKSLREQLRPP
VAEETRARLSKELQAQAARLGDHEDVCGRIYVRGEVQAMLGSTELRRLASHLR
KIKRLKLDADDLQRLAVQAGARGEKELSAIRETLGFLVEQGRVATVSLAQ
QPLQRRAAQGEKRLRARMESGSRTRDLDFVKQVAREVRAKLEEQAOQJRLQAEARQ
ARLKSMEFLVEDMDQWAGHVEKVAQVAVGTSAAPVPSDNH"
61. .114
/gene="APOE"
/note="apolipoprotein E"
115. .1011
/gene="APOE"
/product="apolipoprotein E mature peptide"
267
/gene="APOE"
/note="g in [U. Biol. Chem. 258, 11422-11422 (1983)]. [3],
[4]"
/citation=[2]
/replace="c"
279
/gene="APOE"
/note="g in [U. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
/citation=[4]
/replace="a"
288
/gene="APOE"
/note="g in [U. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
/citation=[4]
/replace="a"
369
/gene="APOE"
/note="in variant"
/citation=[4]
/replace="a"
409
variation

```

```

/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="a"
448
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="c"
538
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="t"
568
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="c"
586
variation
/ gene="APOE"
/ note="in epsilon-2 allele"
/ citation=[3]
/ citation=[5]
/ replace="t"
702
variation
/ gene="APOE"
/ note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
/ citation=[2]
/ replace="g"
783
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="t"
858
variation
/ gene="APOE"
/ note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3] and [4] variant; a in [4] normal,"
/ citation=[4]
/ replace="a"
BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN
Alignment Scores:
Pred. No.:      9,71e-81      Length:      1156
Score:          1585.00      Matches:      317
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             9      Gaps:      0
US-09-827-854-15 (1-317) x HUMAPOE3 (1-1156)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPhenLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      61 ATGAAGGTTCTGGGGCTGGTGGCTGCTACATTCCTGGCAGAGTCCGAGCCAAAGTG 120
QY      21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
      |||||||
Db      121 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAGAC 180
QY      41 GlyGlnArgTrpGlnLeuAlaLeuGlnGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
Db      181 GGCCACGCGCTGGGAGACTGGCAGCTGGCTCTTTGGGATTACCTGGCGCTGGTGCAGACA 240
QY      61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
      |||||||
Db      241 CTGCTGAGCAGGTGAGGAGGAGCTGCTCAGCTCCAGGTCACCCAGGAAGTGAAGGCG 300
QY      81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
      |||||||
Db      301 CTGATGAGCAGAGCCATGAAGGAGTTGAAGGCTTACAAATCGGAAGTGGAGGACAAC 360
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="a"
448
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="c"
538
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="t"
568
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="c"
586
variation
/ gene="APOE"
/ note="in epsilon-2 allele"
/ citation=[3]
/ citation=[5]
/ replace="t"
702
variation
/ gene="APOE"
/ note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
/ citation=[2]
/ replace="g"
783
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="t"
858
variation
/ gene="APOE"
/ note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3] and [4] variant; a in [4] normal,"
/ citation=[4]
/ replace="a"
BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN
Alignment Scores:
Pred. No.:      9,71e-81      Length:      1156
Score:          1585.00      Matches:      317
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             9      Gaps:      0
US-09-827-854-15 (1-317) x HUMAPOE3 (1-1156)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPhenLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      61 ATGAAGGTTCTGGGGCTGGTGGCTGCTACATTCCTGGCAGAGTCCGAGCCAAAGTG 120
QY      21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
      |||||||
Db      121 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAGAC 180
QY      41 GlyGlnArgTrpGlnLeuAlaLeuGlnGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
Db      181 GGCCACGCGCTGGGAGACTGGCAGCTGGCTCTTTGGGATTACCTGGCGCTGGTGCAGACA 240
QY      61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
      |||||||
Db      241 CTGCTGAGCAGGTGAGGAGGAGCTGCTCAGCTCCAGGTCACCCAGGAAGTGAAGGCG 300
QY      81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
      |||||||
Db      301 CTGATGAGCAGAGCCATGAAGGAGTTGAAGGCTTACAAATCGGAAGTGGAGGACAAC 360

```

Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyerdun, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/URL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

source

1..1186
/organism="Homo sapiens"
/db_xref="LocusID:348"
/db_xref="taxon:9606"
/clone="MGC:1571 IMAGE:3355712"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NHL MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
46..999
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAH03557.1"
/db_xref="GI:13097699"
/translation="MAKYMALLVFLAAGCAKVEQVETPEPELNOQTEWQGRN ELAAGRFMDYLRWVOTLSBOVELSSOVQELRALMDTEMKLAKYKSELEOLTP VAETRRARLSDLELOAAQARLGADEMDVCGRLVQYRGVQAMLGSTELRLASHLR KLRRLRLADLDQRLKRLAVQAGAREGERSLIRREGLRIGVGGHRAVATVSLG OPIORAAQWMEERLRARMEGSRTRDLEVRKQVAVAKLEQAOQIRLQAEARO ARLSWFEPLCEVROMAGIVEKQVAAVSTSAAPVSDNH"

CDS

BASE COUNT 248 a 366 c 425 g 147 t
ORIGIN

Alignment Scores:

Pred. No.: 9.97e-81 Length: 1186
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-15 (1-317) x BC003557 (1-1186)

QY 1 MetysValleuTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGTTCTGTGGCTGCGTGTGTCACATTCCTGGCAGATGCCAGGCCAAGCTG 105
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 106 GAGCAAGCGGTGGAGACAGACCCGAGCGCCGACGACACACAGCGAGCGAGAGC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 166 GCGCCAGCGCTGGAACTGGCAGCTGCTTTTGGATTTACCTGCGCTGGGTGCAGACA 225
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 226 CTGCTGTGACAGTGCAGAGAGAGCTCAGCTCCAGGTCACCCAGAGAACTGAGGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 286 CTGATGACGAGACCATGAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAAACACTG 345
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 346 ACCCGGTGGCGAGAGAACCGCGGACGCTGTCCAAAGAGACTCAGGCGGCGAGGCC 405
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnTyrArgGlyGlnVal 140

DB 406 CGCGTGGCGCGACATGAGAGAGCTGTGCGCGCCCTGTGTCAGTACCGCGGAGGCTG 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCAATGCTGGCGACAGACACGAGAGCTGCGGTGCGCTGCTCCCACTGCGC 525
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AAGCTGCTTAAGCGGCTCTCCGCGCATGCGATGACGACAGAGCGGCTGGCACTGAC 585
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 586 CAGCGCGGCGCGCGGAGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 645
QY 201 ProLeuValGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 646 CCGCTGGTGAACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
QY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlnGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 706 CTACAGAGAGCGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCGCGGAGGCG 765
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
DB 766 AGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGG 825
QY 261 LeuGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 826 CTGGAGGAGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCG 885
QY 281 SerTrpPheGluProLeuValGlnLysPheMetGlnArgGlnTrpAlaGlnValGlnValGln 300
DB 886 AGCTGTTGAGACCGCTGCTGGAGAACATGACAGCGCCAGTGCGCGGCGGCTGGTGAAGA 945
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAGCGCTCGCGTGGCAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 996

RESULT 7

E08423 1110 bp RNA linear PAT 29-SEP-1997
LOCUS E08423 DNA coding human apolipoprotein E3.
DEFINITION E08423
ACCESSION E08423.1 GI:2176540
VERSION E08423.1
KEYWORDS JP 1994315392-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1110)
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euthera; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Morimoto, H. and Teranishi, Y.
TITLE METHOD FOR PRODUCING APOLIPOPROTEIN
JOURNAL Patent: JP 1994315392-A 1 15-NOV-1994;
MITSUBISHI KASEI CORP

COMMENT

OS Homo sapiens (human)
PN JP 1994315392-A/1
PD 15-NOV-1994
PF 11-JUN-1985 JP 1994015433
PI MORIMOTO HIRONORI, TERANISHI YUTAKA
PC C12P21/02.C07K13/00.C12N5/10.C12N5/12.C12N5/18.(C12P21/02,
C12N5/10.C12N1:91);
CC strandedness: Double;
CC topology: linear;
FH key Location/Qualifiers
FT source 1..1110
FT sig_peptide 15..68
FT mat_peptide 69..968
FT /product='human apolipoprotein E3'.
FEATURES source location/Qualifiers 1..1110

Db 361 ACCCGGTGGCGAGAGACGGCGCACGGCTGTCCAAAGAGATGCAGGCGGCGAGGCC 420
 Oy 121 ArgLeuGluAlaAspMetGluAspValIcysGluArgLeuValGlnTyrArgGluVal 140
 Db 421 CGCGTGGCGGCGACATGAGAGACTGTGCGCCCGCTGGTGCACTACCGGCGAGGTG 480
 Oy 141 GlnAlaMetLeuGluGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCGCATGCTCGGCCACAGACCGAGAGACTCGGGGTGGCCCTCCACCTGGCCG 540
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 541 AAGCTGCTAAGCGGCTCTCCCGCATGCTCGCAAGATGCTGGCACTGTAC 600
 Oy 181 GlnAlaGluAlaArgGluGluValGluArgGluLeuSerAlaIleArgGluArgLeuGly 200
 Db 601 CAGCGCGGGCGCGCGAGGGCGCGAGCGCGCTCAGCGCCATCGGAGCGCTGGGG 660
 Oy 201 ProLeuValGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGTGGAAACAGGGCGCGGTGGGGCCGCACTGTGGCTCCCTGGCGCGCAGCGC 720
 Oy 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
 Db 721 CTACAGAGCGGGCGCCAGGCTGGGGCGAGCGGCTGCGCGCGATGAGAGATGGGC 780
 Oy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
 Db 781 AGCGCGACCGCGACCGCTGGAGAGGTGAAGAGAGGTGGCGGAGGTGGCGCCCAAG 840
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgLeuLys 280
 Db 841 CTGGAGAGAGAGCGCCACAGATACGCTGCAGAGCGCGGCTTCAGAGCCCGCTCAAG 900
 Oy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGCTTCAGAGCCCTGCTGGAAACATGCAGCCGAGTGGCGGCTGGTGGAAG 960
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGCAGGCTGCGGTGGSCACAGCGCGCCCTGTGTGCCAGACAAATCAC 1011
 RESULT 9
 AX333278 1157 bp DNA linear PAT 09-JAN-2002
 LOCUS Sequence 3787 from Patent W00194629.
 ACCESSION AX333278
 VERSION AX333278.1 GI:18123912
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
 Horrigan,S., Soppet,D.R. and Weaver,Z.
 Cancer gene determination and therapeutic screening using signature
 TITLE
 JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
 AVALON Pharmaceuticals (US)
 FEATURES
 source 1..1157
 location/Qualifiers
 /db_xref="taxon:9606"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.1e-80 Length: 1157
 Score: 1576.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 99.43% Indels: 0

DB: 6 Gaps: 0
 US-09-827-854-15 (1-317) x AX333278 (1-1157)
 Oy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGGTTCGTGTGGTGGCTGGCTGGTGCACATTCCTCGGAGATGCGCAGGCCAAAGTGTG 121
 Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCCACAGACCGAGTGGCAGAGC 181
 Oy 41 GlyGlnArgTrpGluLeuAlaLeuGluValArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 182 GCCCAGCCCTGGGAACCTGGCACTGGGTGGCTTTTGGGATTACTCGCTGGGTGCAGACA 241
 Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 242 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTACCCCAAGAACTGAGGGCG 301
 Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 Db 302 CTGATGGACGAGACCATGAAGATTGAAGGCTTCAATGGAATGGAATGGAAGAACACTGG 361
 Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 362 ACCCGGTAGCGGAGAGACCGCGGCGACGGCTGTCCAAAGAGACTCAGACGGCGCAGGCC 421
 Oy 121 ArgLeuGlnAlaAspMetGluAspValIcysGluArgLeuValGlnTyrArgGluVal 140
 Db 422 CGCGTGGCGGCGACATGAGAGAGCTGTGGCGCCCTGTGTGTCATACCGCGGCGAGGTG 481
 Oy 141 GlnAlaMetLeuGluGlnSerThrGluLeuArgValAlaGlnAlaSerHisLeuArg 160
 Db 482 CAGGCGATGCTCGCGCAGAGCACCGAGAGCTGCGGGTGGCGCTGCCCTCCACCTGGCG 541
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 542 AAGCTGCTAAGCGGCTCTCCCGCATGCCATGACTGCAGAACGGCTGGCACTGTAC 601
 Oy 181 GlnAlaGluAlaArgGluGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGCGCGGGCGCGCGAGGGCGCGAGCGCGCTCAGCGCCATCGCGAGCGGCTGGGG 661
 Oy 201 ProLeuValGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGTGGAAACAGGGCGCGGTGGGGCCGCACTGTGGCTCCCTGGCGCGCAGCGC 721
 Oy 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
 Db 722 CTACAGAGCGGCGCCAGGCTCGGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGG 781
 Oy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGAGCGGTGGCGGAGGTGGCGCCCAAG 841
 Oy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgLeuLys 280
 Db 842 CTGGAGAGAGAGCGCCACAGATACGCTGCAGAGCGCGGCTTCAGAGCCCGCTCAAG 901
 Oy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 902 AGCTGCTTCAGAGCCCTGCTGGAAACATGCAGCCGAGTGGCGGCTGGTGGAAG 961
 Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGCGGTGGSCACAGCGCGCCCTGTGTGCCAGACAAATCAC 1012
 RESULT 10
 AX409597 1157 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 2244 from Patent W00229103.
 DEFINITION AX409597
 ACCESSION

VERSION AX409597.1 GI:21442302
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Alvarès, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2244 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M12529"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 3.1e-80 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: Gaps: 0
US-09-827-854-15 (1-317) x AX409597 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGTTCTGTGGCTGCTGCTGTGGTGCACATTCCTGGCAGAGGCCAGGGG 121
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGAGACCGAGTGGCAGAG 181
QY 41 GylGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACGCTGGGAGTGCACACTGGCTCTTTGGATTACCTCGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAlaGln 80
DB 242 CTGTCTGACAGAGTGCAGAGGAGCTGTCCAGCTCCCAAGATCCCAAGAGTGCAGAGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGln 100
DB 302 CTGATGAGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAATCGAGAGCAACTG 361
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCCGCTAGCGGAGAGAGCGGGCGACGCTGCCAAGAGCTGCAGAGCGGCGAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyLysVal 140
DB 422 CGGCTGGGGGACATGAAGAGCTGTGCGGCGCTGTGCTGCAATCGCGCGAGG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATCTCGCCAGACAGACCGAGAGCTGGGGTGGCTCCCTCCCACTGGCCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCTGAACGCGCTCTCCGCGATCCGATGACCTGCAGAAAGCGCTGGGAGTGC 601
QY 181 GlnAlaGlyAlaArgGlnGlnLysAlaGlyLysSerAlaLeuArgGlnArgLeuGly 200
DB 602 CAGGCGGGGGCCGAGAGGCGCCGAGCGGCTTCAGGCGCATCCGAGAGCGCTGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGGTGAACAGAGCGCGCTGGCGGCGCCACTGTGGCTCTCTGGCGGAGCCG 721

QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuAlaArgMetGluMetGly 240
DB 722 CTACAGAGAGCGGCGCCAGGCTGGGGGAGCGGCTCGCGCGGATGAGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 782 AGTGGACCCCGGACCGCTCGGAGAGGTGAAGAGACAGAGTGGAGGAGCGCGCAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnAlaLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGAGAGGCGCCAGCAGATAGCTGACAGCGGAGGCTTCAGGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnThrAlaGlyLeuValGluLys 300
DB 902 AGCTGTTGAGACCCCTGTGGTGAAGACATGACGCCCACTGAGCGCGCTGTGGAGAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspSerHis 317
DB 962 GTGCAGGCTGCGCTGGGACACAGCGCGCCCTGTGCCAGACCAATCAC 1012
RESULT 11
LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5472858.
ACCESSION 115975
VERSION 115975.1 GI:1250883
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Attle, A.D., Gretch, D.G., Sturley, S.L. and Beckage, N.E.
TITLE Production of recombinant proteins in insect larvae
JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
FEATURES
Location/Qualifiers
1..1157
/organism="unknown"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 3.1e-80 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: Gaps: 0
US-09-827-854-15 (1-317) x 115975 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGTTCTGTGGCTGCTGCTGTGGTGCACATTCCTGGCAGAGGCCAGGGG 121
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGAGACCGAGTGGCAGAGC 181
QY 41 GylGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACGCTGGGAGTGCACACTGGCTCTTTGGATTACCTGCGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAlaGln 80
DB 242 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCACCTCCCAAGTCCCAAGAGTGCAGAGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGln 100
DB 302 CTGATGAGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAATCGAGAGCAACTG 361
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
DB 362 ACCCGGTTAGCGGAGAGAGCGGGCGACGCTGCCAAGAGCTGCAGAGCGGCGAGGCC 421

QY	121	ArgLeuGluGlyAlaAspMetGluAspValGlySerGlyArgLeuValGlnTyrArgGlyGluVal	140
QY	121	ArgLeuGluGlyAlaAspMetGluAspValGlySerGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	422	CGGCTGGGGCGGCAATGAGAGACGTGTGGCGGCCCTGGTGCACATACCGCGGCGAGGTG	481
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGluLeuAspValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCCATGCTGGCGCAGAGCACCCAGAGAGCTGGCGGGTGGCGCTGCTCCACCTGGGC	541
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	542	AAGCTGCCTAAGCGGCTCTCTCCGCATCTCCCATGACCTGCAGAAACGCCCTGGCAGTGTAC	601
QY	181	GlnAlaGlyAlaArgGluGlyValaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	602	CAGGCCGGGGCCCGCGAGAGGCCCGCAGAGCGGGCTTCACGCGCCATCCGCGAGCGCTGGGG	661
QY	201	ProLeuValGluGlnGlyArgValArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro	220
Db	662	CCCCGTGGTGAACAGAGGCGCGCTGGGGGCGGCACATGTGTGGCTCTGCGCGGACGCG	721
QY	221	LeuGlnGluArgAlaGlnAlaThrPolyGluArgLeuArgAlaArgMetGluGluMetGly	240
Db	722	CTACAGAGAGCGGGCCAGAGCCCTGGGGCGAGCGGCTGCCGCGCATGAGAGATGGGCG	781
QY	241	SerArgTyrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys	260
Db	782	ACTCGGACCCCGGACCGCTTGAGCAGAGGTGAAGAGAGGTGGCGGAGGTGGCGCCCAAG	841
QY	261	LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
Db	842	CTGGAGGAGAGAGGCCAGAGATACGCTGCAGGCGCGAGGCTTCCAGGCGCGCTCAAG	901
QY	281	SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuValGluLys	300
Db	902	ACCTGCTTCGACCCCTCGTGTGGAACATGACAGCCGACGTGGCGGCTGTGTGAGAAG	961
QY	301	ValGlnAlaAlaValAlaGlyThrSerLeuAlaProValProSerAspAsnHis	317
Db	962	GTGCGAGCTCGCGTGGGCAACGAGCCGCCCTGTGTGCCAGGACAAATCAC	1012
RESULT	12		
HUMAPOE	HUMAPOE	1157 bp	mrna linear PRI 08-AUG-1995
LOCUS	Human apolipoprotein E, mRNA, complete cds.		
DEFINITION	M12529		
ACCESSION	M12529.1	GI:178848	
VERSION			
KEYWORDS	apolipoprotein.		
SOURCE	Homo sapiens (clone: pHAEl112,178,813). male 57-year old liver cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1157) McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.		
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant		
JOURNAL	J. Biol. Chem. 259 (10), 6498-6504 (1984)		
MEDLINE	84212473		
PUBMED	6327682		
FEATURES			
source	Location/Qualifiers		
	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:3606"		
	/map="19q13.2"		
	/clone="pHAEl112,178,813]."		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="57-year old"		
	1..1157		
	/gene="APOE"		
	<1..1157		
mrna			

CDS	/gene="APOE" /note="apoE mRNA; G00-119-691" 62..1015 /gene="APOE" /note="precursor" /codon_start=1 /product="apolipoprotein E" /protein_id="AA859518.1" /db_xref="GI:178849" /db_xref="GDB:G00-119-691" /translation="MKVLAALLVTFLAGCOAKVEAVETEPERPELROQTEMQ ELAGRPMDYLRMWQTESEVOVELSSQYTOELRALMDETMEKLKAYSELE VAESTERRLSKELOTAQRILGADMEDCGRLIYRGEOAMIGOSTEELRVRI KLRRRLRDPDDLOKRAIVQAARBEAEGSLTERGRPLVEDGSRYAAVR OPLERAQAWGERIRARRAMEMSGSTRDRIDEVEQAEVRAKULEEQAOQIRLD ARIKSWEPLVEDMQROMAGLVVERQAAVGTSAAPVSDNH" 62..115
sig_peptide	/gene="APOE" /note="G00-119-691" 116..1012
mat_peptide	/gene="APOE" /product="apolipoprotein E" /note="G00-119-691"
variation	370 /gene="APOE" /note="a in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	410 /gene="APOE" /note="a in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	569 /gene="APOE" /note="a in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	784 /gene="APOE" /note="t in PHAE[112,178]; c in PHAE813; G00-119-691" /replace="c"
variation	859 /gene="APOE" /note="g in PHAE[112,178]; a in PHAE813; G00-119-691" /replace="a"
BASE COUNT	212 a 370 c 426 g 149 t
ORIGIN	30 bp upstream of bstNI site.
Alignment Scores:	
Pred. No.:	3,1e-80 Length: 1157
Score:	1576.00 Matches: 315
Percent Similarity:	99.37% Conservative: 0
Best local Similarity:	99.37% Mismatches: 2
Query Match:	99.43% Indels: 0
DB:	Gaps: 0
US-09-827-854-15 (1-317) x HUMAPOE (1-1157)	
Oy	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Dd	62 ATGAAGCTTCTGTGGCGCTGCTGTCATCTCCGCGAGGATGCCAGGCCAAGTG 121
Oy	21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Dd	122 GAGCAAGCGGTGGAGACAGACC GGAGCCGAGCTGC GCCAGCAGACCGAGTGCAAGC 181
Oy	41 GlyGlnArgGTPGluLeuAlaLeuGlyValArgPheTrpAspTryLeuAgtTrpValGlnThr 60
Dd	182 GCCCAGCGCTGGGAATGGCACTGGGTCTTTGGATTACCTGGCCCTGGGTCCAACA 241
Oy	61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Dd	242 CTGTGTGACGACGTCCAGGAGGACTGCTCAAGCTCCCAAGTCAACCAAGAATGAGGCG 301
Oy	81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrlYssSerGluLeuGlnGlnLeu 100

Db 841 CTGGAGGACGAGCCGACGATACGCTGTGAGGCCGCTTCCAGGCCCTCAAG 900

QY 281 SerTrrPhegluProluValgluAspmetGlnArgIntrrPalaglYleuValgluYrs 300
|||||
Db 901 ACCTGTTGAGCCCTGGTGGAGACATGACGCCCATGTGGCCGCTGGTGGAGAAG 960

QY 301 ValGlnAlaAlaValglYThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 961 GTGACAGGCTGGGAGCACGAGCCGCCCTGTGGCCAGCAGACATCAC 1011

RESULT 14

MEAPOE 1178 bp mRNA linear PRI 31-MAR-1995

LOCUS DEFINITION Monkey mRNA for apolipoprotein E.

ACCESSION X13887

VERSION X13887.1 GI:38054

KEYWORDS apolipoprotein; apolipoprotein E.

SOURCE Macaca fascicularis.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca; Cercopithecoidea; Macaca.

REFERENCE 1 (bases 1 to 1178)
Marotti,K.R.
Direct Submission
Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and
Melchior,G.W.
Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300

TITLE Data kindly reviewed (20-Mar-1989) by Marotti K.R.

JOURNAL Location/Qualifiers

MEDLINE 1. 1178
PUBMED /organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_id="lambda gll1"
83. 1036
/codon_start=1
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/translation="MKYLMALVTFLGCAQKVEQPEPEPELROAQEGSGQPM
ELALGRFDYLRWYOTLSEVOEELSPVOBELTTMDETMKELKAYKSLLEOLSP
VAETTRARLSKEIOAAOARIGADMEDVRSRYOVRSEVOAMLGSTETRLARLSHR
KLKRLRLDADLQRLAVYQAGRBEGAEVSAIREFLGLVDEQGRVRAVTSLAS
QPLDERALGLERLARMEMGSRRLDELDEKQVAVRAKLEEQAOQISLQAEAFQ
ARLSWEPELVEDMROWAGLYEKVQAAVASTAPVPIDNH"

CDS
83. 1036
/product="proapolipoprotein E (AA -18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/translation="MKYLMALVTFLGCAQKVEQPEPEPELROAQEGSGQPM
ELALGRFDYLRWYOTLSEVOEELSPVOBELTTMDETMKELKAYKSLLEOLSP
VAETTRARLSKEIOAAOARIGADMEDVRSRYOVRSEVOAMLGSTETRLARLSHR
KLKRLRLDADLQRLAVYQAGRBEGAEVSAIREFLGLVDEQGRVRAVTSLAS
QPLDERALGLERLARMEMGSRRLDELDEKQVAVRAKLEEQAOQISLQAEAFQ
ARLSWEPELVEDMROWAGLYEKVQAAVASTAPVPIDNH"

sig_peptide 137. 1033
mat_peptide /product="apolipoprotein E (AA 1 - 299)"
misc_feature 1154. 1159
/note="polya signal"

BASE COUNT 217 a 373 c 426 g 162 t

ORIGIN

Alignment Scores:

Pred. No.: 3 13e-73 Length: 1178

Score: 1451.00 Matches: 296

Percent Similarity: 94.01% Conservatave: 2

Best Local Similarity: 93.38% Mismatches: 19

Query Match: 91.55% Indels: 0

DB: 9 Gaps: 0

US-09-827-854-15 (1-317) x MEAPOE (1-1178)

QY 1 MetLysValIeuTrpAlaIleuValThrPheLeuAlaGlyCysGlnAlaIalysVal 20

Db 83 ATGAAGTTCTGTGGGCTGCTGGTGGTGCACATTCCTGGAGATGCCAGGCCAAGTG 142
|||||

QY 21 GluGlnAlaValglYThrGluProgluIleuArgGlnGluThrGluTrpGlnSer 40
|||||

Db 143 GAGCAACCGGTGGAGCCAGACAGACGAGACCGAGCTTGGCAGACAGCTGAGGCCAAGC 202
|||||

QY 41 GlyGlnArgTrpGluIleuAlaLeuGluArgPheTrpAspTyrLeuArgTrrValGlnThr 60
|||||

Db 203 GGCAGACCCCTGGAGACTGGACATGGGTGCTTTTGGAGTAACTGCGGTGGTGGAGACA 262
|||||

QY 61 LeuSerGluGlnValgluGluIleuLeuSerSerGlnValThrGlnIleuArgAla 80
|||||

Db 263 CTGTCTGAGCAGGTGACAGAGAGAGAGCTGCACGCCCCAGAGTCACCCAGAGAACTGACAGACG 322
|||||

QY 81 LeuMetAspGluThrMetLysGluLeuYsaIalYrLysSerGluLeuGluIleuGlnLeu 100
|||||

Db 323 CTGATGGAGACCATGATGAAGGCTTGAAGGCCCTCAAAATCGGAACCTGAGAGACAGCTG 382
|||||

QY 101 ThrProValAlaGluIleuThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||

Db 383 AGCCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGACTGCAGCGCGCCAGGCC 442
|||||

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrrArgGlyGluVal 140
|||||

Db 443 CGGTGGGTGGCGACATGAGAGACGCGGCGACCGCGCTGTGTGACAGACGAGCGAGGTG 502
|||||

QY 141 GlnAlaMetLeuGlyGlnSerThrGluIleuArgValArgLeuAlaSerHisLeuArg 160
|||||

Db 503 CAGGCCATGCTGGGCGAGGAGTACCGAGAGCTGCGGCGCGCTGCTCCACCTGCGCC 562
|||||

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||

Db 563 AAGCTGCGCAAGCGGCTCTCCGCGATGCTGATGACTGCACAGAGCGCTGGCAGTGTAT 622
|||||

QY 181 GlnAlaGlyAlaArgGluGluIleuArgLysGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||

Db 623 CAGGCGGGGGCGCGGAGGCGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGGA 682
|||||

QY 201 ProLeuValgluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyIlePro 220
|||||

Db 683 CCCCTGGGTGAGACAGGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 742
|||||

QY 221 LeuGlnIleuArgAlaGlnAlaThrPglYgluArgLeuArgAlaArgMetGluIleuGly 240
|||||

Db 743 CTTCAGGAGCGGCGGCGGAGGCTTGGGTGAGCGGCTTGGCGGAGATGAGAGATGGGC 802
|||||

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||

Db 803 AGCGGAGCGCGGCGGAGCGGCTGTGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 862
|||||

QY 261 LeuGluGluGlnAlaGlnGlnIleuArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||

Db 863 CTGGAGGAACAGCGCCAGACAGATAGCTGTGAGCGCGGAGGCGCTTCCAGGCGCGCTCAAG 922
|||||

QY 281 SerTrrPhegluProluValgluAspmetGlnArgIntrrPalaglYleuValgluYrs 300
|||||

Db 923 ACCTGTTGAGCCCTGGTGGAGATATGACGCCCATGTGGCTGGGTGGAGAAG 982
|||||

QY 301 ValGlnAlaAlaValglYThrSerAlaAlaProValProSerAspAsnHis 317
|||||

Db 983 GTGACAGGCTGGGAGCACGAGCCGCCCTGTGGCCAGCAGACATCAC 1033
|||||

RESULT 15

AF261279 5491 bp DNA linear PRI 27-OCT-2000

LOCUS DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.

ACCESSION AF261279

VERSION AF261279.1 GI:11034800

KEYWORDS Homo sapiens.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS	1 (bases 1 to 5491) Nickerson, D.A., Taylor, S.L., Fullerton, S.M., Weiss, K.M., Clark, A.G., Stengard, J.H., Saloma, V., Boerwinkle, E. and Sing, C.F.
TITLE	Sequence diversity and large-scale typing of SNPs in the human apolipoprotein E gene
JOURNAL	Apollipoprotein E gene
MEDLINE	Genome Res. 10 (10), 1532-1545 (2000)
PUBMED	20499366 11042151
REFERENCE	2 (bases 1 to 5491) Nickerson, D.A.
AUTHORS	Direct Submission
TITLE	Submitted (27-MR-2000) Department of Molecular Biotechnology, University of Washington, Box 357730, Seattle, WA 98195, USA
JOURNAL	Location/Qualifiers
FEATURES	1..5491 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="19" /map="19q13.2" /3..>108 /note="putative" /rpt_family="MIR" /rpt_type="dispersed 73 /frequency="0.01" /replace="t" /replace="t" /note="putative" /rpt_family="MIR" /rpt_type="dispersed 308 /frequency="0.01" /replace="t" /replace="t" /note="putative" /rpt_type="tandem <320..>339 <320..>339 /note="putative" /rpt_type="tandem <340..>637 /note="putative" /rpt_family="Alu" /rpt_type="dispersed 471 /frequency="0.01" /replace="g" /replace="g" 345 /frequency="0.01" /replace="t" /replace="t" 560 /frequency="0.22" /replace="t" /replace="t" 624 /frequency="0.07" /replace="c" /replace="c" <638..>718 /note="putative" /rpt_type="tandem 832 /frequency="0.45" /replace="t" /replace="t" join(1060..1094,1855..1920,3013..3205,3786..4645) /product="apolipoprotein-E" 1163 /frequency="0.35" /replace="c" 1522 /frequency="0.01" /replace="a" 1575 /frequency="0.01" /replace="t" /replace="t" join(1878..1920,3013..3205,3786..4503) /note="APOE" /codon_start=1

variation	1998	/product="apolipoprotein-E" /protein_id="AAG27089.1" /db_xref="GI:11034801" /translation="MKVLAALVTFLAGCAKVAQVETPEPEPLRQQTQMGSCRR ELAGREWDYLRWOTLSEQVQELLSSQVQELRALMDETMKELKAKSELBQUTP VAEERFRLRSKELQAAQRLGADMEDVCGRLVQYRGVQAMIGSTELRYRLASHLR KLRLRLDADDDLOKRLRLVYQAGAREGARELSAIREHIGPLVEGGRATVGSILAG QPLDRAQAGEBRLRARMEEGMSRTROLDLVKRDQVAVRKLLEBQAQQRILQAEAFQ ARLSWFEPLVEDMQROMAGLVERVQAQVGSAAPIVPSDNN" 1998
repeat_region		/frequency="0.10" /replace="a" <2124..>2435 /note="putative" /rpt_family="Alu" /rpt_type-dispersed 2440
variation		/frequency="0.21" /replace="a" <2569..>2848 /note="putative" /rpt_family="Alu" /rpt_type-dispersed 2907
variation		/frequency="0.01" /replace="g" 3106 /frequency="0.01" /replace="c" <3472..>3588 /note="putative" /rpt_family="Alu" /rpt_type-dispersed 3673
repeat_region		/frequency="0.01" /replace="g" 3701..3702 /frequency="0.01" /replace="ct" 3937
variation		/frequency="0.14" /replace="c" 4036 /frequency="0.01" /replace="t" 4075
variation		/frequency="0.07" /replace="t" <4755..>5056 /note="putative" /rpt_family="Alu" /rpt_type-dispersed 5065..>5476 /note="putative" /rpt_family="Alu" /rpt_type-dispersed 5229..5230 /frequency="0.03" /replace="g" 5229
variation		/frequency="0.07" /replace="t" 5229..5230 /frequency="0.40" /replace="g" 5230
variation		/frequency="0.13" /replace="a" 5361 /frequency="0.06" /replace="c" /repeat="c" BASE COUNT 1029 a 1603 g 1196 t ORIGIN

Alignment Scores:

Piped. No.: 1.73e-69 Length: 5491
 Score: 1396.50 Matches: 304
 Percent Similarity: 60.808 Conservative: 0
 Best Local Similarity: 60.808 Mismatches: 2
 Query Match: 88.11% Indels: 194
 DB: 9 Gaps: 1

us-09-827-854-15 (1-317) x AF261279 (1-5491)

QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
 Db 3003 TTCCACACAGGATGCGCAGGCCCAAGGTGGAGCAAGCGGTGGAGACAGAGCCGAGCCGAG 3062
 QY 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPhe 51
 Db 3063 CTGCGCCAGCAGACCGAGTGGAGCGGCGCGCTGGGAACCTGGCACTGGGTGCTTT 3122
 QY 52 TrpAspTyrLeuArgTyrValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSer 71
 Db 3123 TGGGATTACCTGCGCTGGGTGAGACACTGCTGTGAGCAGAGTGCAGAGAGAGCTGCTCAGC 3182
 QY 72 SerGlnValThrGlnGluLeu----- 78
 Db 3183 TCCAGGTGACCCAGAGAACTGAGGTGATGCCCATCTGGCCCTTGACCCCTCTGGTG 3242
 QY 78 ----- 78
 Db 3243 GCGGCGTATACCTCCCGCAGGTCCAGGTTTCATTTCTGCCCTGTGCTAAGTCTGGGGGG 3302
 QY 78 ----- 78
 Db 3303 COTGGGTCTGCTGCTGCTCTAGCTTCCCTTCCATTTCTGACTCTGCTTACCTTACCTCTC 3362
 QY 78 ----- 78
 Db 3363 TGGAACTCTCTCTACGCTTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 3422
 QY 78 ----- 78
 Db 3423 CGTCTGCT 3482
 QY 78 ----- 78
 Db 3483 CTCACGTGTGCGCCAGCGCTGCTTGAACCTTCTGGGCTCAAGCATCTCCGCTCGG 3542
 QY 78 ----- 78
 Db 3543 COTCCAAAGTGTGGATAGAGCATGAGCACCCTTGGCCGCTCTAGCTCTCTCT 3602
 QY 78 ----- 78
 Db 3603 TCGTCTGCTCTGCTCTGCTCTGCACTCTCTCTGCACTCTCTCTCTCTCTCTCTCTCTG 3662
 QY 78 ----- 78
 Db 3663 GCCTCTGCCCGTCT 3722
 QY 78 ----- 78
 Db 3723 CCCCATTCCAGCCCTTCTCCCGCCTCTCCACTGTGCGACACCCTCCGCTCTCTGCGCG 3782
 QY 79 -ArgAlaLeuMetAspGluThrMetLeuLysGluLeuLysAlaTyrLysSerGluLeuGluGlu 98
 Db 3783 CAGGCGCTGATGAGCGAGACCATGAGAGAGTGAAGGCTTACAAATCGAAGCTGGAGGA 3842
 QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
 Db 3843 ACAACTGACCCCGGTGGGAGAGACGCGGCGACAGGCTGTCCAAAGAGAGCTGCAGCGCG 3902
 QY 118 aGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlu 138

Db 3903 GCAGGCCGCTGGCGCGGACATGAGAGACGTGTGCGCGCCGCTGTGCTACGTACCGCGG 3962
 QY 128 YGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHis 158
 Db 3963 CGAGGTGAGGCGATGCTCTGGCGCAGAGCACCGAGAGAGCTGCGGGTGGCTCTGCTCCCA 4022
 QY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAla 178
 Db 4023 CCGCGCAAGCGCGCTAAGCGGCTCTCTCGCATCCCATGACCTGCAGAAACGCGCTGCG 4082
 QY 178 aValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgLysLeuSerAlaAlaArgGluArg 198
 Db 4083 AGTATACAGCGCGCGCGCGCGAGGCGCGCGAGCGCGCGCTCAGCGCATCCGCGAGCG 4142
 QY 198 gLeuGlyProLeuValGluGlnGlyArgValArgAlaArgAlaThrValGlySerLeuAlaGlu 218
 Db 4143 CCGGGGCGCGCTGTGGAACAGAGCGCGCTGCGGCGCGCGCATCTGTGGCTCTCTGCGCG 4202
 QY 218 YGlnProLeuGlnGluArgAlaGlnAlaArgPglYgluArgLeuArgAlaArgMetGluGlu 238
 Db 4203 CCAGCGCTACAGAGAGCG 4262
 QY 238 uMetGlySerArgTyrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 258
 Db 4263 GATGGCGAGCGGACCG 4322
 QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArg 278
 Db 4323 CGCCAGGCTGGAGAGAGCG 4382
 QY 278 gLeuLysSerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuVal 298
 Db 4383 COTCAAGAGCTGTGCTGAGCGCGCTGTGGAAGACATGACAGCGCGCGCGCGCGCGCGCTGT 4442
 QY 298 IGlulysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 4443 GGAGAGGTGAGAGCTGCGGTGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4500

Search completed: March 14, 2003, 17:32:12
 Job time : 2314.23 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds

(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589

Sequence: 1 MKVLMAALLVFLAGCQAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/runat.11032003.101610.27486/app-query.fasta.1.3576
-DB=GenEmbl -QWRT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=-1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pct -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn.1.1.13008.@runat.11032003.101610.27486 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: GenEmbl:*
2: gb_ba:*
3: gb_hg:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1589	100.0	1156	6	BD004277 Apo E hum
2	1577	99.2	1110	6	E00359 E00359 CDNA coding
3	1577	99.2	1110	6	E00823 E00823 DNA sequence
4	1577	99.2	1147	6	AX302545 AX302545 Sequence
5	1577	99.2	1156	6	BD004278 Apo E hum
6	1577	99.2	1156	6	BD004278 Apo E hum
7	1577	99.2	1186	9	BC003557 Homo sapi
8	1570	98.8	1110	6	E08423 E08423 DNA coding
9	1568	98.7	1157	6	AX333278 AX333278 Sequence
10	1568	98.7	1157	6	AX409597 AX409597 Sequence
11	1568	98.7	1157	6	I15975 I15975 Sequence 1
12	1568	98.7	1157	6	HUMAPOE HUMAPOE
13	1565	98.5	1176	6	BD004279 BD004279 Apo E hum
14	1443	90.8	1158	9	MEAPOE MEAPOE
15	1388.5	87.4	5491	6	AX358722 AX358722 Sequence
16	1388.5	87.4	41907	9	AF050154 AF050154 Homo sapi
17	1388.5	87.4	107567	9	AC011481 AC011481 Homo sapi
18	1388.5	86.3	5515	9	HUMAPOE4 HUMAPOE4
19	1371	86.3	5515	9	AF261280 AF261280 Pan trogl
20	1342	84.5	5413	9	BARAPOE BARAPOE
21	1359	79.2	4762	2	AC021988 AC021988 Homo sapi
22	1250	78.7	1138	4	AP030380 AP030380 Tupai gl
23	1250	78.7	1138	4	RABAPOLP RABAPOLP
24	1162	75.1	1060	6	AX384545 AX384545 Sequence
25	1155.5	72.7	965	6	BRAPOLPE BRAPOLPE
26	1155.5	72.7	1108	4	AX384541 AX384541 Sequence
27	1155.5	72.7	5617	6	AX384541 AX384541 Sequence
28	1155.5	72.7	6026	6	AX384539 AX384539 Sequence
29	1143	71.9	718	9	AF20049753 AF20049753 Pan trogl
30	1141	71.8	1045	10	MUSAPOE MUSAPOE
31	1141	71.8	1104	10	BC028816 BC028816 Mus muscu
32	1139	71.7	718	9	AF2005052 Gorilla g
33	1139	71.7	718	9	AF2005052 Gorilla g
34	1137.5	71.6	1154	4	AF2005053 AF2005053 Pongo pyg
35	1135	71.4	718	9	BRAPOLPE BRAPOLPE
36	1135	71.4	1122	4	SSAPOE SSAPOE
37	1130	71.1	1126	6	AR164342 AR164342 Sequence
38	1130	71.1	1126	6	AR205885 AR205885 Sequence
39	1086	68.3	951	10	S76779 S76779 Mus muscu
40	1064	67.0	959	10	MUSAPOE MUSAPOE
41	1048.5	66.0	1069	10	RATAPOLP RATAPOLP
42	1021	64.3	228698	2	AC127479 AC127479 Mus muscu
43	1021	64.3	237653	2	AC073760 AC073760 Mus muscu
44	995	62.6	4856	10	MUSAPE MUSAPE
45	994	62.6	4267	6	AR164387 AR164387 Sequence

RESULT 1

ALIGNMENTS

LOCUS	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Apo E humanized mammal.				
ACCESSION	BD004277				
VERSION	BD004277.1 GI:18632238				
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.				
TITLE	Apo E humanized mammal				
JOURNAL	Patent: JP 2001017028-A 1 23-JAN-2001;				
COMMENT	MITSUBISHI CHEMICAL CORP OS Homo sapiens (human) PN JP 2001017028-A/1 PD 23-JAN-2001 PF 28-APR-2000 JP 2000128919 PR				
FEATURES	<p>FT SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC C12N15/09//C07K14/775, PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,C12N5/00,C12R1:91) CC</p>				
source	EH	key	Location/Qualifiers		
	FT	CDS	1..1156 (61)..(1011).		
BASE COUNT	208 a	367 c	432 g	149 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	2.6e-83	length:	1156		
Score:	1589.00	Matches:	317		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-16 (1-317) x BD004277 (1-1156)					
QY	1	MettysValleuTPPALaLaLeuDeuValThrPheLeuAglIyCysGlnAlaLysVal	20		
Db	61	ATGAAGGTTCTGTGGGCTGCGTTCATCTCGCAGAGATGCCAGGCAAGGTG	120		
QY	21	GlUGlnAlaValGluThrGluProGluProGluLeuAurGlnGlnThrGluTrpGlnSer	40		
Db	121	GAGCAAGGGGTGGAGACAGACCAGCCAGCCAGCGCCAGCAGACCCAGAGTGGCAGAGC	180		
QY	41	GLyGlnAaGTTPGluLeuAlaLeuGluYArpPheTrpAspPYrLeuAurGTrpValGlnThr	60		
Db	181	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTTGGATTACCTGCCTGGGTGCAGACA	240		
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValTrhGlnGluLeuAurGlnA	80		
Db	241	CTGTCTGACGAGGTGCAGAGAGAGTGTCTACGCTCCAGGTACACCCAGAACTAGAGGCG	300		
QY	81	LeuMetAspGluTrhMetLysGluLeuLysAlaTrLysSerGluLeuGluGluGlnLeu	100		
Db	301	CTGATGCGACGAGACCATGAAGAGATTGAAGGCTTCAACAATCGGAATCGAGAACACTG	360		
QY	101	ThrProValAglGluGluTrhArgAlaArgLeuSerLysGluLeuGlnAlaAglAla	120		
Db	361	ACCCGGTGGGGGAGAGACACGGGGCAGGGTGTCCAGAGAGCTGCAGGCGCCAGAGCC	420		
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnTrpArgGlyVal	140		
Db	421	CGGCTGGGGCGGAGCATGGAGAGACGTGTGGCGCCCGCTGGTGCAGTACCGGCGGAGGTG	480		
QY	141	GlnAlaMetLeuGlnGlnSerThrGluGluLeuAurGValAArgLeuAlaSerHisLeuAurG	160		

D	b		481	CAGGCATCTCGGCCAGAGCACACGAGACTGGGGTCCGCTCGCCTCCACCTGGCC	540
O	y		161	LysLeuArlYArSArgLeuArlgAspaLaasPspLeuGlLnLysCysLeuAlaValTr	180
D	b		541	AAGCTCGTAAGGGGTCTCTCCGCGATGCGCATGACCTCAAGAAGTGCTGGCAGTGC	600
O	y		181	GlnAlaGlyAlaArgGluunlYalagluAuarqglyLeuSerAlaIleArgGluarGleuLy	200
D	b		601	CAGGGCGGGGGCCCCGACAGGGCCCGAGCGGCGCTTCAGGCCATCCGACGCGCTGGG	660
O	y		201	ProLeuValGlunGlYArGValArVAlaAlaThrValGIySerLeuAlaGlyInPro	220
D	b		661	CCCCTGTGTGAACAGGGGCCGCTGCGGGCGCCCACTGTGTGGCTCCCTGGCCGACC	720
O	y		221	LeuGlunGluaRgaLaGlnAlaTrppGlyGuArGleuArGaLargMetGluGluMetGly	240
D	b		721	CACAAGAGACGGGGCCAGGCTGTGGGGCGAGCGGTGCGCGCGATGAGAGATGGGC	780
O	y		241	SerArgThrArASparGleuaspGluValLlysGluGlnValAlaGluValArGaLays	260
D	b		781	ACCCGACCCCGGACCGCTGTGACGAGGtGAAGAGACAGTGGCGGAGGTGGCGCCAAG	840
O	y		261	LeuGlunGlunAlaGlnGlnIleArgLeuGlnAlaGluAlaPhelGlnAlaArgLeuLys	280
D	b		841	CTGGAGGAGACAGGCCAGCAGATAGCCTGTGACGGCCGAGGCCCTTCAGGCCCGCTCAAG	900
O	y		281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300
D	b		901	ACTGGTTGACGCCCTGTGTGGAAGCATGACAGCGCACTGGCGCGGGCTGTGGAGAG	960
O	y		301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
D	b		961	GTGCAGAGCTGCGGTGGGACACAGCGCCGCTGTGCCACGACAAATCAC	1011
RESULT 2					
E	00359	E00359	1110 bp	RNA	linear
L	OCUS	CDNA coding human apolipoprotein E3.			PAT 29-SEP-1997
D	E00359	E00359			
A	VERSION	E00359.1 GI:2168646			
K	KEYWORDS	JP 1985118189-A/1.			
S	SOURCE	Homo sapiens.			
O	ORGANISM	Homo sapiens			
R	REFERENCE	Emura,Yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
A	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
T	TITLE	Teranishi Y., Takamatsu,N., Matsui,Y., Kimura,M. and Ikeda,Y.			
J	JOURNAL	DNA FRAGMENT			
C	COMMENT	Patent: JP 1985118189-A 1 25-JUN-1985;			
		MITSUBISHI CHEM IND LTD			
		OS human			
		PN JP 1985118189-A/1			
		PD 25-JUN-1985			
		PF 29-NOV-1983 JP 1983224980			
		PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI			
		KIMURA MASAKO,			
		PI IKEDA YASUKO			
		PC C12N15/00.C07H21/04//C12P21/00;			
		CC strandedness: Double;			
		CC topology: Linear;			
		CC hypothetical: No;			
		CC anti-sense: No;			
		CC *source: tissue=liverr;			
		FH key Location/Qualifiers			
		FT CDS	15..968		
		FT sig_peptide	15..68	/product='apolipoprotein E3 precursor' FT	
		FT mat_peptide	69..965	/product='apolipoprotein E3 signal peptide' FT	
		FT Location/Qualifiers			
		FT			

source	1. .1110	/organism="Homo sapiens"	
BASE COUNT	198 a 353 c 416 g 143 t	/db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:	1.22e-82	Length:	1110
Pred. No.:	1577.00	Matches:	316
Score:	99.68%	Conservative:	0
Percent Similarity:	99.68%	Mismatches:	1
Best Local Similarity:	99.24%	Indels:	0
Query Match:	6	Gaps:	0
DB:			
US-09-827-854-16 (1-317) x E00359 (1-1110)			
QY	1	MethylvalleutrpAlaalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
DB	15	ATGAAGTTCTGTGGCTCGTGTGCTGATTCATCTCTGGCAGGATGCCAGGCCAAGGTG	74
QY	21	GIUGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnThr	40
DB	75	GACCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGAGACCGAGTGGCAGAC	134
QY	41	GIYGLnArgTTPGlnLeuAlaLeuGlyArgPheTTPSPtyrLeuArgTTPValGlnThr	60
DB	135	GGCCACGCGTGGGAATGGACACTGGCTCGCTTTGGGATTACTGCGCTGGGTGCAGACA	194
QY	61	LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla	80
DB	195	CTGTCTGACAGGTGACAGAGAGGAGCTGCTGACCTCCAGATCCAGGACCTGAGGCGG	254
QY	81	LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnGlnGlnGlnGlnGln	100
DB	255	CTGATGACAGACACCATGAGAGAGTGTGAGAGCTTACAAATCCGAACTGGAGAGACTG	314
QY	101	ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnGlnGlnGlnGlnGln	120
DB	315	ACCCGCGTGGGAGAGAGAGCGGGCGGCTGTCCAGAGAGCTGCAGCGGGCGGCGGC	374
QY	121	ArgLeuGlnAlaAspMetLysPvalLysGlyArgLeuValGlnThrArgGlyGlnVal	140
DB	375	CGCTGTGGGCGGACATGAGAGACGTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTG	434
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	160
DB	435	CAGGCGATGCTGGCGCAGAGACACCGAGAGCTGGGCTGCGCTGCTGCCACCTGGCG	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180
DB	495	AAGCTCGTAAGGGGCTCTCCGCGATGGCGATGACCTGCAGAGAGCGCTGGCAGGTAC	554
QY	181	GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	555	CAGGCGGGGCGCGCAGAGGCGCGCGAGCGCGCTCAGCGCATCCGACAGCGCTGGGG	614
QY	201	ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	220
DB	615	CCCTGTGTGGAACAGGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	674
QY	221	LeuGlnGlnArgAlaGlnAlaTTPGlyGlnArgLeuArgAlaArgMetGlnGlnGlnGly	240
DB	675	CTACAGAGACGGGCGCAGGCTGGGGCGAGCGGCTGCGCGCGGAGTGGAGAGAGTGGG	734
QY	241	SetArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys	260
DB	735	AGCGGAGCCCGGACCGCTGCGAGAGGTGAAGAGACAGTGGCGGAGGTGGCGGCAAG	794
QY	261	LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	280
DB	795	CTGGAGAGACAGGCCACAGATAGCTGCTGAGCGCGGAGAGGCTTCCAGGCGCGCTCAAG	854
QY	281	SetTTPPheGlnProLeuValGlnAspMetGlnArgGlnTTPAlaGlnLeuValGlnLys	300

Db	855	ACCTGTTGAGACCCCTGTGTGGAAGACATGACAGCCCACTGGGCGGCTGTGTGAGAGAG	914
QY	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspSntHis	317
Db	915	GTGCAGGCTGCGCGTGGGACACAGCGCGCGCTGTGTGCCAGGACATACAC	965
RESULT 3			
LOCUS	E00823	1110 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal peptide.		
ACCESSION	E00823		
VERSION	E00823.1	GI:2169084	
KEYWORDS	JP 1986096997-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1110)		
AUTHORS	Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.		
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN		
JOURNAL	Patent: JP 1986096997-A 1 15-MAY-1986;		
COMMENT	MITSUBISHI CHEM IND LTD		
OS	Human (Homo sapiens)		
PN	JP 1986096997-A/1		
PD	15-MAY-1986		
PF	16-OCT-1984 JP 1984216987		
PI	TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO		
PC	C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12R1:19),		
PC	(C12N15/00,		
PC	C12R1:19);		
CC	strandedness: Double;		
CC	topology: Linear;		
CC	hypotheical: No;		
CC	anti-sense: No;		
CC	*source: tissue_type=Liver; location/Qualifiers		
FT	key		
FT	3'UTR	1..14	
FT	sig_peptide	15..68	
FT		/product='human apolipoprotein E signal	FT
FT	CDS	69..968	
FT		/product='human apolipoprotein E' FT	3'UTR
FEATURES			
source	969..1110. location/Qualifiers		
	1..1110		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	198 a 353 c 416 g 143 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.22e-82	Length:	1110
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-16 (1-317) x E00823 (1-1110)			
QY	1	MethylValleutrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	15	ATGAAGTTCTGTGGCTCGTGTGCTGATTCATCTCTGGCAGGATGCCAGGCCAAGGTG	74
QY	21	GIUGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnThr	40
Db	75	GACCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGAGACCGAGTGGCAGAC	134
QY	41	GIYGLnArgTTPGlnLeuAlaLeuGlyArgPheTTPSPtyrLeuArgTTPValGlnThr	60
Db	135	GGCCACGCGTGGGAATGGACACTGGCTCGCTTTGGGATTACTGCGCTGGGTGCAGACA	194

BASE COUNT	210 a	365 c	425 g	147 t	ORIGIN
Alignment Scores:					
Pred. No.:	1,26e-82	1,577.00	Length: 1147		
Score:	1577.00	99.68%	Matches: 316		
Percent Similarity:	99.68%	99.68%	Conservative: 0		
Best Local Similarity:	99.68%	99.68%	Mismatches: 1		
Query Match:	99.24%		Indels: 0		
DB:	6		Gaps: 0		
US-09-827-854-16 (1-317) x AX302545 (1-1147)					
Oy	1	MettysValleuTrippaAaLaLeuLeuValThrpheLeuAlaGlyCysGlnAlaLysVal	20		
Db	46	ATGAAGGTTCTGTGGGCTGGCTTCTGTGTACATTCTCGGAGGATCCAGGCCAAGTG	105		
Oy	21	GlunGlnAlaValGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer	40		
Db	106	GAGCAAGCGGTGGAGACAGAGACCAGGAGCCGACGTCGGCCAGCAGAGCGAGTGCCAAAGC	165		
Oy	41	GlyGlnAArgTrpGluLeuAlaLeuGlyAArgpHeRpbAspTryLeuAArgTrpValGlnThr	60		
Db	166	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGATTACCTGGCCCTGGGTGCACACA	225		
Oy	61	LeuSerGlnGlnValGlnGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuAArgAla	80		
Db	226	CTGTCTGAGCAGGTGGACAGAGAGACTGTGTACGCTCCAGGTCAACCCAGAACTGAGGGCG	285		
Oy	81	LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGlnGlnLeu	100		
Db	286	CTGAATGGACGAGACCATTGAAGAGATTGAAGCCCTCAAAATCGGAACGTGAGGAACAACCTG	345		
Oy	101	ThrProValAlaGlnGluThrAArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
Db	346	ACCCCGGTGGCGGAGAGAGACCAGCGGCAAGCTGTCTCCAAAGAGCTGCAGCGCGCCAGGCC	405		
Oy	121	ArgLeuGlnAlaAspMetGluAspValCysGlyAArgLeuValGlnTryAArgGlyVal	140		
Db	406	CGCGTGGCGCGGACATGAGAGAGCTGTGCGCGCCCTGTGGTCAAGTACCAGCGGAGAGTG	465		
Oy	141	GlnAlaMetLeuGlnGlnSerThrGlnGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160		
Db	466	CAGGCCATGCTCGGCGAAGACACCGAGAGAGCTGGCGGGTGCCTCGCCTCCACCTTGGC	523		
Oy	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180		
Db	526	AAGCTGCTAAGCGGCTCTCCGCGATGCGCATGACCTGCAGAAAGCCCTGGCAGTGTAC	585		
Oy	181	GlnAlaGlyAlaAArgGlnGlyAlaGlnAArgLysLeuSerAlaLeaArgGluAArgLeuGly	200		
Db	586	CAGCGCGGGGCGCCGAGAGGCGCGCGACGCGCTTCAGAGCCGATCCCGAGCGCTGGGG	645		
Oy	201	ProLeuValGlnGlnGlnGlyAArgValAArgAlaAlaThrValGlySerLeuAlaGlnPro	220		
Db	646	CCCCGTGGTGAACAGGGCGCGGTGCGGGCCGCCACTGTGGGCTCTCGCGCGGCCAGCCG	705		
Oy	221	LeuGlnGlnAArgAlaGlnAlaAArgPglyLysLeuAArgAlaAArgMetGlnGlnMetGly	240		
Db	706	CTACAGGAGCGGCGCCAGAGGCTGTGGGCGAGCGCTGCGCGCGCGCGATGAGAGAGATGGCG	765		
Oy	241	SerArgTrpAArgAspAArgLeuAspGluValLysGlnGlnValAlaGluValAArgAlaLys	260		
Db	766	AGCCGGAGCCCGCGACCGCTGTGAGAGAGGTGAAGAGAGAGGTGGGGAGGTGGCGCCAAG	825		
Oy	261	LeuGlnGlnGlnAlaGlnGlnGlnAlaAArgLeuGlnAlaGlnAlaAArgGlnAlaAArgLeuLys	280		
Db	826	CTGGAGGAGCAGGCCCAAGACATACGCTGTGAGGCCAGAGGCTTCCAGAGGCCCGGCTCAAG	885		
Oy	281	SerTrpPheGluProLeuValGluAspMetGlnAArgGlnTrpAlaGlyLeuValGluLys	300		
Db	886	AGCTGGTTTCGAGCCCTGTGTGGAAGACATCACAGGCCACAGTGTGGCGCGGTGTGGAGAAG	945		

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 946 GTGCAGGCTGCGGTGGGACACGAGCGCCCTGTGTGCCAGAGCAATCAC 996
 RESULT 5
 LOCUS BD004278 1156 bp DNA linear PAT 31-JAN-2002
 DEFINITION Apo E humanized mammal.
 ACCESSION BD004278
 VERSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Fujita,S., Hamanaka,H.,
 TITLE Apo E humanized mammal
 JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PE 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
 A01667/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10,PC
 C12N5/09//C07K14/775,
 PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC
 FEATURES
 source location/Qualifiers
 FT Key CDS location/Qualifiers
 (61)..(1011).
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.27e-82 Length: 1156
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: Gaps: 0
 US-09-827-854-16 (1-317) x BD004278 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGCGTGTGTCACATTCCTGCGCAGATGCCAGGCCAAGGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGGGGTGAGACAGAGCGGAGCGGAGCTGCGCGCAGACACGAGAGTGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCCAGCGCTGGAGCACTGGCTGCTTGGATTCCTGCGCTGGGTGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTGTGACGAGTCCAGAGAGAGCTGCTCAGCTCCAGGTCAACCAGGAACCTGAGGCGC 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
 DB 301 CTGATGAGCAGACCATGAGAGAGTTGAAGGCTTCAATCGAATCGAGACTGAGAGAACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGAGCTCAGCGCGCGCAGCGC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 421 CCGCTGGGCGCGGACATGAGAGACCTGTGGCGCCCTGTGTACATACCGCGGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGCGCCAGAGCACCGAGGAGCTGCGGTGCCCTGCCCTCCACCTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 541 AAGCTGCTTAAGCGGCTCTCCGCGATGCGGTGACCTGCACAAAGCGCTGTGCATGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyValGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCCATCGCGAGGCGCTGGG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaTrpValGlySerLeuAlaGlnPro 220
 DB 661 CCCCTGTGGAAACAGAGCGCGCTGCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 221 LeuGlnGluArgAlaGlnAlaLeuTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
 DB 721 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGCGGAGTGAAGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 AGCGGAGACCGCGGACCGCTGTGACAGGTGAAGAGAGCAGTGGCGAGTGGCGCCCAAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGAGAGCGCCACAGATACGCTGACGCGCGAGCGCTTCAGAGCGCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 AGCTGTGTGACCGCTGCTGTGAAGACATGACGCCCACTGGGCGCGGTGTGTGAAGAG 960
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTGCGGTGGGACACGAGCGCCCTGTGTGCCAGAGCAATCAC 1011
 RESULT 6
 LOCUS HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 DEFINITION Homo sapiens preapoliipoprotein E (APOE) mRNA, complete cds.
 ACCESSION K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apoliipoprotein; apoliipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,
 TITLE Identification and DNA sequence of a human apoliipoprotein E cDNA clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 REFERENCE
 AUTHORS Wallis,S.C., Rogne,S., Gyll,L., Markham,A., Edge,M., Woods,D.,
 TITLE The isolation of cDNA clones for human apoliipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 REFERENCE
 AUTHORS Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.
 TITLE Synthesis, intracellular processing, and signal peptide of human apoliipoprotein E


```

                                and [4] variant: a in [4] 'normal'
                                /citation=[4]
                                /replace="a"
BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN

Alignment Scores:
Pred. No.:      1,27e-82      Length:      1156
Score:          1577.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:      99.24%      Indels:      0
DB:              Gaps:      0

US-09-827-854-16 (1-317) x HUMAPOE3 (1-1156)
OY      1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
        |||||||
DB      61 ATGAAGGTTCTGTGGCGCTGCTGTGTGCATTCCTGGCAGATGCCAGGCCAAGGTG 120
        |||||||

OY      21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGlnThrGlnThr 40
        |||||||
DB      121 GAGCAAGCGGTGAGACACAGAGCCGAGCCGAGCTGCCAGCAGACGAGTGGCAGAGC 180
        |||||||

OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
        |||||||
DB      181 GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGGGATTTACTGCGTGGGTGCAGACA 240
        |||||||

OY      61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
        |||||||
DB      241 CTGTGTGAGCAGGTGACAGAGAGCTGCTCACCTCCAGAGTCCAGAGAACCTGAGCGCG 300
        |||||||

OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnGln 100
        |||||||
DB      301 CTGATGAGCAGACCATGACATGAGAGCTTGAAGGCTTACAAATCGGAATGGAGACACACTG 360
        |||||||

OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
        |||||||
DB      361 ACCCGGCGGGGAGAGAGCGGGGACGCGCTGCCAAGAGACTGACAGCGCGCGCAGCGC 420
        |||||||

OY      121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
        |||||||
DB      421 CGCGTGGGGCGGACATGAGAGACCTGTGCGCGCCCTGTGTCACGTACCGCGCGCAGAGTG 480
        |||||||

OY      141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
        |||||||
DB      481 CAGGCGCATGCTGGCCAGAGACACCGAGGAGCTGCGGGTGGCTGCTCCACCTGGCGC 540
        |||||||

OY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
        |||||||
DB      541 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACCTGCAGAACCGCTGGCAGTGTAC 600
        |||||||

OY      181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
        |||||||
DB      601 CAGGCGCGGGCGCCGAGGAGCGCCGAGCGCGCTCAGCGCATCCGCGAGCGCTGGGG 660
        |||||||

OY      201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
        |||||||
DB      661 CCCCTGGTGGAGACAGGCGCGCTGCGCGCTGAGTGGGTCTCCCTGGCGCGCAGCGG 720
        |||||||

OY      221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
        |||||||
DB      721 CTACAGAGAGCGGGCGCCAGGCGCTGGGCGCGCTGCGCGCGGATGGAGAGATGGGC 780
        |||||||

OY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
        |||||||
DB      781 AGCCGAGACCCGCGACCGCTGGACAGGTGAAGGAGACGAGTGGCGGAGTGGCGCCCAAG 840
        |||||||

OY      261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
        |||||||
DB      841 CTGGAGGAGACAGGCGCCAGCAGATACGCTTCCAGGCGCGGCTCCAGG 900
        |||||||

OY      281 SerTrpPheGluProLeuValGluLysPheGlnArgGlnTrpAlaGlyLeuValGluLys 300
        |||||||

```

```

DB      901 ACCTGCTCGACCCCTGCTGAGACATGACAGCCCGCATGGCGCGCTGCTGAGAGAG 960
        |||||||
OY      301 ValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
        |||||||
DB      961 GTTCAGGCTGCCGTGGGACACAGCGCGCCCTGTGTGCCAGACATATCAG 1011
        |||||||

```

```

RESULT 7
LOCUS      BC003557
DEFINITION Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA,
complete cds.
ACCESSION BC003557.1 GI:13097698
VERSION   BC003557
KEYWORDS  MGC.
SOURCE    Homo sapiens.
ORGANISM  Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1186)
REFERENCE
AUTHORS   Strausberg, R.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

REMARK
COMMENT
Contact: MGC help desk
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BC DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Beta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pavan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schelin, Diane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuyerdunyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

CDS

```

1..1186
/organism="Homo sapiens"
/db_xref="locusID:348"
/db_xref="taxon:9606"
/clone="MGC:1571 IMAGE:3355712"
/tissue="Eye, retinoblastoma"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
46..999
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAH03557.1"
/db_xref="GI:13097698"
/translation="MKVLMALVTFLAGCAQKAEQAVETPEPELRLNQETWESGGRN
ELAGRWFDYLRWVOTLSEQVDELSSQVDELRLMDETMKELAKYKSLSEQLTP
VAEETRLARLKEILOAAOARLQADMEDVCGRLVQYRGVEQVALGOSTEELRRLASLR
KLRRLRLDADDLQRLAVYQAGAREGERGLSAIRELIGLVQGGRRVATVSLG
QPLQERQAWGERLRARMEGSRTRIDRDEKQVAVAKLEQAOQIRLQAEARQ
ARKLSWEPLVEDQROMAGIYVEQAAVGSAPVPSDNI"
BASE COUNT      248 a      366 c      425 g      147 t
ORIGIN

```


Db 315 ACCCGGTGGCGAGAGACGGCGGCGACGGCTGTCCAGAGCTGCAGGCGCGCAGGCC 374
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 375 CGGCTGGGGCGGACATGGAGGACGTGTGCGGCGCGCTGTGTGACAGTACCGCGCGAGGTG 434
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGGCGCATCTCGGCGCAGACACCGAGAGCTGCGGCTGCGCTGCCCTCCACCTGCCG 494
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 495 AAGCTCGGTAACCGGCTCTCCGCGATGCCATGACATGACAAAGCGCTGGCAGGTGAC 554
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 555 CAGGCGGGGGCGCGCGAGGGCGCGCGCGCGCTTCAGCGCATCCGCGAGCGCTGGGGG 614
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 615 CCCGTGTGGAAACAGGGCGCGGTGGCGCGCGCTGTGTGCTGCTGCCGCGCGCAGCGC 674
Qy 221 LeuGlnGluArgAlaGlnAlaThrPheGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 675 CTACAGAGAGCGGGCGCGAGCGCTGGGGCGCGCGCTGGCGCGATGGAGAGATGGGC 734
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 735 AGCGGAGACCGCGCGCGCTGGAGCGGTGAGAGAGAGAGAGTGGCGAGGTGGCGCGCAAG 794
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 795 CTGGAGAGACAGCGCCAGAGATACGCTCGCAGGCGAGGCGCTTCAGGCGCGCGCTCAAG 854
Qy 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPheGlyLeuValGluLys 300
Db 855 AACTGGTTGACCGCTGTGTGAGACATGACAGCGCGCATGGCGCGGTGTGTGAGAGA 914
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 915 GTGCAGGCTGCGGTGGGACACGCGCGCGCTGTGTGCCAGGACATCAC 965

RESULT 9
AX333278 1157 bp DNA linear PAT 09-JAN-2002
LOCUS AX333278
DEFINITION Sequence 3787 from Patent WO0194629.
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, R.C., Edner, R., Endress, G.,
Horrikan, S., Soppet, D.R., and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
FEATURES
source
1. 1157
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 4.2e-82 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0

DB: 6 Gaps: 0
US-09-827-854-16 (1-317) x AX333278 (1-1157)
Qy 1 MetLysValLeuTrrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGAGCTGCGCTTGGTGTGACATTCCTGTGGAGATGCGCAGGCG 121
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrrPheGlnSer 40
Db 122 GAGCAAGCGGTGGAGACAGCGGAGCGCGAGCTGCGCAGACAGACAGAGCGAGAGG 181
Qy 41 GlyGlnArgTrrPheGluLeuAlaLeuGlyArgPheTrrPheTrrPheArgTrrPheGlnThr 60
Db 182 GGCAGCGCTGGAGACTGGCAGCTGGGTGCTTTGGATTTACTGCGCTGGGTGAGAGA 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTGTAGACAGGTGACAGAGAGAGCTGTACAGCTCCCAAGTCAACCAAGACTGAGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrLysSerGluLeuGlnGluGlnLeu 100
Db 302 CTGATGGAGACAGCATGAAGGATTTGAAGGCTTCAAAATTCGAATCGAGAGACAACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGAGAGCGGCGCAGCGCTGTCCAGAGAGCTGCAGCGCGCAGGCG 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGGGCGGACATGAGAGAGAGCTGTGCGCGCGCTGTGTGCTGTGTGCTGTGCT 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATCTGCGCGCAGAGACCGCAGAGAGCTGGGGGTGGCTGTGCTCCACTGTGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 542 AAGCTGGCGTAACCGGCTCTCGCGCATCCGATGACCTGCAGACAGCGCTGCGCATGTAC 601
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGCGCGCGAGGCGCGCAGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCGTGTGGAAACAGGGCGCGGTGGCGCGCGCTGTGTGCTGCTGTGCGCGCAGCGC 721
Qy 221 LeuGlnGluArgAlaGlnAlaThrPheGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 722 CTACAGAGAGCGGCGCGAGGCGCTGGGGCGAGCGCGCTGCGCGCGGATGGAGAGATGGGG 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGAGACCGCGCGCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCGGCGCGAGATACGCTGCAGGCGCGAGGCTTCAGAGCGCGCTCAAG 901
Qy 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrrPheGlyLeuValGluLys 300
Db 902 AAGCTGTTCAGAGCGCTGTGTGAACATGACAGCGCGCATGGCGCGGTGTGTGAGAGA 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGACACGCGCGCGCTGTGTGCCAGGACATCAC 1012

RESULT 10
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS AX409597
DEFINITION Sequence 2244 from Patent WO0229103.
ACCESSION AX409597

VERSION	AX409597.1	GI:21442302
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Alvarares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.	
AUTHORS	1	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 2244 11-Apr-2002;	
FEATURES	GENE LOGIC INC (US)	
source	Location/Qualifiers	
	1..1157	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/note="EMBL/Genbank Accession No. M12529"	
BASE COUNT	212 a	370 c 426 g 149 t
ORIGIN		
Alignment Scores:		
Pred. No.:	4.2e-82	1157
Score:	1568.00	Matches: 314
Percent Similarity:	99.05%	Conservative: 0
Best Local Similarity:	99.05%	Mismatches: 3
Query Match:	98.68%	Indels: 0
DB:	6	gaps: 0
US-09-827-854-16 (1-317) x AX409597 (1-1157)		
QY	1	MettysValleuTPRPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB	62	ATGAAGGTTCTGTGGCGCTGCTGCTGTCATTCCTCGAGAGATCCAGGCCAAGTG 121
QY	21	GlGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGInSer 40
DB	122	GAGCAAGCGGTGGAGACAGAGCCGGAGCCACGCTCGCCAGCAGACCGAGTGGCAAGC 181
QY	41	GlyGlnArgTPGluLeuAlaLeuGlyAlaPheThrPaspArgLeuArgTPValGlnThr 60
DB	182	GGCCAGCGCTGGAGACTGGCACTGGGTGGCTTTGGGATTACCTGGCGTGGGTGCAGACA 241
QY	61	LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB	242	CTGTCTGACGAGTGGACAGAGAGACTGCTGCTGCCACCAAGTCACCCAAAGATGAGGCGC 301
QY	81	LeuMetAspGluThrMetIysGluLeuLysAlaIlyIysSerGluLeuGlnGlnLeu 100
DB	302	CTGATGGACGAGACCATGTAAGGATTGAAGCCCTTCAAAATTCGGAACGGAGAACCACTG 361
QY	101	ThrProValAlaGlnGluThrArgAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAla 120
DB	362	ACCCCGGTAGCGGAGAGAACCGGGGACAGGCTGTTCCAAGAGCTGCAGACGGCCAGGCC 421
QY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB	422	CGGCTGGCGCGGACATGAGAGAGCTGGCGCGCCCTGGGCGAGTACC CGCGGAGGTG 481
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB	482	CAGGCGATGCTGGCCCAAGACACCGAGAGAGCTGGGGGTGGCGCTCCGCTCCACCTGCC 541
QY	161	LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIlyScySleuAlaValTyr 180
DB	542	AAGTGCTTAAGCGGCTCTCTCCGGAATCCGATGACTCTGCAGAAAGCCCTGGCAGTGTAC 601
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB	602	CAGCGCGGGGCCCGCGAGGGCGCGAGCGGGGCTTCAGCCCATCCCGAGCCGCTGGGG 661
QY	201	ProLeuValGlnGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlnPro 220
DB	662	CCCTGGTGGAGACAGGGCGGTGGGGCGCCACTGTGGGCTTCCCTGGCGGGGCACCG 721

QY	221	LeungJingLiuAraAlaGlnAlaATrrpolylgluaArgleuAraAlaArgmetCluclumetgly	240
Db	722	CTACAGAGAGCGGGCCAGGCCCTGTGGGCGAGCGCGTGCCTGGCATGAGAGATGTGGC	781
QY	241	SerArqThrArAspArpdeuaspgluValLySgJugInVaIalagluValargrAlaLyS	260
Db	782	AOTCGAGCCCCGGACCGCCTGACCAAGGTGAAGAGACAGTGCCGAGGTGGCGCCCAAG	841
QY	261	LeuJuglugInlaGlngInlleaArgleugInlaGluAlaPheGlnAlaArgrleuLyS	280
Db	842	CTGGAGAGACAGCCCGACAGATACGCTTCAGAGCCGAGGCTTCAGGCCCGCTCAAG	901
QY	281	SerTrpphegluProleuValglunuspmetGlnaArgGIntPrAlaGlyleuValgluLyS	300
Db	902	AACTGGTTTCGAGCCCCTGGTGGAAACAATGACAGCCCACTAGTGGCGCGGTGGTAGAGA	961
QY	301	ValGlnAlaAlaValaGlyThrSerAlaAlaProVerAspAsnHis	317
Db	962	GTCGAGCGCTGGCTGGCACCAAGCCCGCCCTGTGTGCCAGGACAAATCAC	1012
RESULT	11		
LOCUS	I15975	1157 bp	DNA Linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	I15975		
VERSION	I15975.1 GI:1250883		
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.E. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNAL	Patent: US 5472858-A 1 05-DEC-1995;		
FEATURES	Location/Qualifiers		
Source	1..1157 /organism="unknown"		
BASE COUNT	212 a 370 c 426 g 149 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4.2e-82	Length:	1157
Score:	1568.00	Matches:	314
Percent Similarity:	99.05%	Conservative:	0
Best Local Similarity:	99.05%	Mismatches:	3
Query Match:	98.68%	Indels:	0
DB:	6 Gaps:	0	
US-09-827-854-16 (1-317) x I15975 (1-1157)			
QY	1	MettysValLeutrpAlaAlaLeuValThrPheLeuAlaGlyCySGlnAlaAlaLySVal	20
Db	62	ATGAAGGCTTCTGTGGGCTGCTGCTGTCACATTCTCGGACAGATGCCAGGCCAAGGTG	121
QY	21	GJugInlaAlaValaGluThrGluProGluProGluLeuArGIngInlThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCAGCCAGACAGCAGAGTGCAGAGC	181
QY	41	GlyGlnArqTrpGluLeuAlaLeuGlnArqPhetrPaapTryLeuArqTrpValGlnThr	60
Db	182	GGCCAGCCCTGGGAACCTGGCACTGGGTGGCTTTGGATTAACCTGCGCTGGGTGCAGACA	241
QY	61	LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArGrAla	80
Db	242	CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAAGTCCAACATCCACCAAGAACTAGAGGCG	301
QY	81	LeuMetaspGluThrMettysGluLeuLySaIaTryLySerGluLeuGluGluGlnLeu	100
Db	302	CTGATGGACAGACCATGAGAGATTGAAGGCTTCAAATCGGAACCTGAGAGAACACTG	361
QY	101	ThrProValAlaGluGluThrArqAlaArqLeuSerLySgJugInVaIalagluAla	120
Db	362	ACCCTGGTGGAGAGAACCGGGCAGCGGTGTCTCAAGAGATCTCAACAGCGCGCAGGCC	421

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 422 CGGCTGGGGCGGACATGAGAGACTGTGGCGCCGCTGGTGCATGACCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCGATGCTGCGGCAGAGACACCGAGAGCTGGCGGTGGCTGCTCCACACTGGCG 541
 QY 161 LysLeuArgGlyArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 542 AAGCTGCTAAGCGGCTCTCCGCGATCCGATGACCTGCAGAAAGCGCTGGCACTGTAC 601
 QY 181 GlnAlaGlyAlaArgGlyGlnGlyValGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGCGCGGGGGCGCGAGGGGGCGCGAGCGCGCTCCAGCGCCATCCGAGAGCGCTGGGG 661
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGGTGAACAGGGCGCGCTGGCGGCGCACTGTGGCTCTCCGCGGCGAGCGG 721
 QY 221 LeuGlnGluArgAlaGlnAlaIleTyrGlyLysArgLeuArgAlaArgMetGluGluMetGly 240
 Db 722 CTACAGAGAGCGGGCCAGGCTGGGGCGAGCGGCTGGCGGCGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 Db 782 AGTCGAGACCGCGACCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAG 901
 QY 281 SerTrpPheGlnProLeuValGluAspMetGlnArgGlnTyrPheGlnGlyLeuValGluLys 300
 Db 902 AGCTGTGTCGAGCCCTGGTGAAGACATGACAGCCGAGCGGCGGCTGTGTGAGAG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGCAGGCTGGCGGTGGGACACAGCCGCCCTGTGTGCCAGAGAGAGATCAC 1012
 RESULT 12
 HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 LOCUS Human apolipoprotein E mRNA, complete cds.
 DEFINITION M12529
 ACCESSION M12529.1 GI:178848
 VERSION M12529.1 GI:178848
 KEYWORDS apolipoprotein.
 SOURCE Homo sapiens (clone: PHAE112,178,813.) male 57-year old liver CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source location/Qualifiers
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE112,178,813".
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 mRNA

/gene="APOE"
 /note="APOE mRNA; G00-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AA859518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:G00-119-691"
 /translation="MKVLMALLVTFLAGCQAKVEQAVETPEPELRQQTWMSGQRM ELAIGRFWDYLRWVOTLSEVOBELSSOYTOELRALMDETMKLAKVSELEQOLRP VAEETRARLSKELOTAORLIGADMEYCGRLVOYRGVQAMLCOSTEELRYRLASHR KLRKRLRDDDDLOKRLAVYQAGAREAGELISIRREILGPLVCGGVRRAVYSLAG OPIQROANGCERLRARMEGSRTPRRLDEVRQVAFVRAKLEBQAOQRLQAEARQ ARLSKRPFLVEDHQROMAGLIVERVQAAVGTSAAPVPSDNI"
 62..115
 sig_peptide
 /gene="APOE"
 /note="G00-119-691"
 116..1012
 mat_peptide
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"
 370
 variation
 /gene="APOE"
 /note="a in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 410
 variation
 /gene="APOE"
 /note="a in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 569
 variation
 /gene="APOE"
 /note="c in PHAE112,178]; g in PHAE813; G00-119-691"
 /replace="g"
 784
 variation
 /gene="APOE"
 /note="t in PHAE112,178]; c in PHAE813; G00-119-691"
 /replace="c"
 859
 variation
 /gene="APOE"
 /note="g in PHAE112,178]; a in PHAE813; G00-119-691"
 /replace="a"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.
 Alignment Scores:
 Pred. No.: 4.2e-82 Length: 1157
 Score: 1568.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.68% Indels: 0
 DB: 9 Gaps: 0
 US-09-827-854-16 (1-317) x HUMAPOE (1-1157)
 QY 1 MetLysValLeuTrpAlaIleLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 62 ATGAAGTTCTGTGGGCTGGCTGTGCTCAATTCCTGGCAAGGATGCGAGAGAGGTG 121
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 122 GAGCAAGCGGTGGAGACAGACCGGAGCGCGAGCGCGAGAGAGAGAGAGAGAGAGAG 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 182 GGCACGCGCTGGAACTGGCACTGGCTGCTTTGGGATTCCTGCGCGGTGGTGAAGCA 241
 QY 61 LeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
 Db 242 CTGCTGAGCAGGTGACAG 301
 QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100

|||||
Db 302 CTGATGAGACGACCTGAAAGAGTTGAAAGCCTTACAAATTCGAACTGGAGGACACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGCGTAGCGAGAGACAGCGGGCAGCGCTGTCCAGAGAGCTGCAGACGGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetLysPvalCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGCGCGGACATGAGAGCGTGTGCGGCGCTGTGACAGTACCGCGCGGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATCTCGGCCAGAGACACCGAGAGCTGGGGTGGCTCGCTGCCCTCCACCTGGCC 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 542 AAGCTGCGTAAGGGCTCTCTCCGCGATCCCGATGACCTGCAGAGAGCGCTGGAGTGA 601
Qy 181 GlnAlaGlyAlaArgGluGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCGCGAGGCGCGCGGCTTCAGCGCCATCCGAGCGGCTGGGG 661
Qy 201 ProLeuValGluGluGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGTGTGAACAGGGCGCGGCTGCGGCCGCACTGTGGCTCCCTGGCGGCGAGCCG 721
Qy 221 LeuGlnGluArgAlaGlnAlaThrPglLysGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGG 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 ACTGCGGACCGCGGAGCGCTGTGACGAGGTGAAGAGACAGCTGGCGGAGGTGGCGCCAG 841
Qy 261 LeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGACAGGCGCCAGAGATACGCTGCAGGCGGAGCCCTCCAGAGCCCGCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 ACTGCTGTGAGACCCCTGTGTGGAAGCATGACAGCGCAATGGGCGCGCTGTGGAGAA 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCGCTGGCGCACGACGCGCGCCCTGTGCGCCAGGACATATCAC 1012
RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279 Apo E humanized mammal.
DEFINITION
ACCESSION BD004279.1 GI:18632240
VERSION JP 2001017028-A/3.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
A0167/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10,PC
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

Alignment Scores:
Pred. No.: 6,24e-82 Length: 1156
Score: 1565.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.49% Indels: 0
DB: Gaps: 0
US-09-827-854-16 (1-317) x BD004279 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCTGCTGTGTGTCACATTCCTGGCAGATGCCAGGCCAAAGTG 120
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAAGACAGAGCGGAGCCGAGCTGGCCACAGACGAGTGGCAGAGC 180
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCACGCGTGGGAAGTGGCACTGGGTGCTTTTGGAATTAATCTGCGCGGTGGCAGACA 240
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGACAGAGTGCAGAGAGAGCTGTGCACCTCCAGGTCAACCAAGAACTGAGAG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db 301 CTGATGAGACAGACATGAAGAGGTGAAGGCTTCAAAATCGAACTGAGAGAACACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGCGTGGCGAGAGAGAGCGGGGACCGCTGTCCAAAGAGCTCAGCGGCGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetLysPvalCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGCTGGCGGCGCCCTGTGTGCAATACCGGCGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCCAGAGACACCGAGAGCTGGGGTGGCTGCTGCCCTCCACCTGGCG 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGCTCTCTCCGCGATGCCATGTACCTGCAGAACCGCTGCACTGAT 600
Qy 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 601 CAGGCGGGGGCGCGAGAGGCGCCAGAGCGGCGCTCAGCGCCATCCGAGAGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGAACAGGGCGCGGCTGGCGGCGCACTGTGGGCTCCTGGCGGCGAGCCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaThrPglLysGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGACCGCGGAGCGCTGTGAGAGGTGAAGAGAGCTGGCGGAGTGGCGGCCCAAG 840
Qy 261 LeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||

Dn	841	CTGGAGGACGAGCCCGAAGATACGCCTGCCAGGCGAGCCCTTCCAGGCCCGCCTCAAG	900
Qy	281	SerTrpPhegluProleuValAlguaspmetGlInrGlnTrpAlaglyLeuValglulys	300
Dn	901	AGCTGTTCGAGCCCTGCTGGGAAGACATGCACGCCACTGGGCGCGCTGTGGAGAAG	960
Qy	301	ValGlnAlaIalValgIyThrsrAlaIalProvalProserAspashis	317
Dn	961	GTCAGAGCTGCGTGAGGACCAGCGCCCGCTGTGCCAGCACATCATC	1011
RESULT 14			
LOCUS	MFAPOE	1178 bp	mRNA linear PRI 31-MAR-1995
DEFINITION	Monkey mRNA for apolipoprotein E.		
ACCESSION	X13887		
VERSION	X13887.1 GI:38054		
KEYWORDS	Apolipoproteins; Apolipoprotein E.		
SOURCE	Macaca fascicularis.		
ORGANISM	Macaca fascicularis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;		
TITLE	Cercopitheidae; Macaca.		
JOURNAL	Marotti, K.R.		
AUTHORS	1 (bases 1 to 1178)		
TITLE	Direct Submission		
JOURNAL	Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,		
AUTHORS	301 Henrietta Street, Kalamazoo, MI 49008		
TITLE	2 (bases 1 to 1178)		
JOURNAL	Marotti, K.R., Whitted, B.E., Castle, C.K., Polites, H.G. and		
AUTHORS	Melchior, G.W.		
TITLE	Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA		
JOURNAL	Nucleic Acids Res. 17 (4), 1778 (1989)		
PUBMED	89160349		
COMMENT	2922300		
FEATURES	Data kindly reviewed (20-Mar-1989) by Marotti K.R.		
source	Location/Qualifiers		
	1..1178		
	/organism="Macaca fascicularis"		
	/db_xref="taxon:9541"		
	/clone="1"		
	/tissue_type="liver"		
	/clone_lib="lambda gt11"		
	83..1036		
	/codon_start=1		
	/product="proapolipoprotein E (AA -18 to 299)"		
	/protein_id="CAA32092.1"		
	/db_xref="GI:38055"		
	/translation="MKVIMALIVTFLLGCAKVEQPYEPTEPELKOQAEGSGGOW		
	ETALGAEFSLRWOTLSBOVEELISFOVTQETTLMDETMKELEAKYSLEQDLSIP		
	VAEETFRWDESKELQAQARLGADMEVDVRSLEYOAMLGSTEELERLASHSR		
	KIKRKLRIDADDLQRLAVYQAGAREGEERGVSAIRELIGLVQGRVRAATVSLS		
	OPIOEAFALGERLARAMEGMSPTRDDEVKEDVAVRALKLEDAQQLSIQAEARQ		
	ARKSMFEELVEDMQROMAGLIEVKYAAGASTAPVIDNH"		
	83..136		
	sig_peptide		
	mat_peptide		
	137..1033		
	/product="apolipoprotein E (AA 1 - 299)"		
	misc_feature		
	1154..1159		
	/note="polya signal"		
	BASE COUNT	217 a 373 c 426 g 162 t	
ORIGIN			
Alignment Scores:			
Pred. No.:	6..54e-75	Length:	1178
Score:	1443..00	Matches:	295
Percent Similarity:	93.69%	Conservative:	2
Best Local Similarity:	93.06%	Mismatches:	20
Query Match:	90.81%	Indels:	0
DB:	9	Gaps:	0
US-09-827-854-16 (1-317) x MFAPOE (1-1178)			
1 MetlyValleutrpAlalaaleuValThrPhleuAlaIglyCysGlnAlaIylsVal			
20			

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
TITLE Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
JOURNAL Sequence diversity and large-scale typing of SNPs in the human
MEDLINE apolipoprotein E gene
PUBMED Genome Res. 10 (10), 1532-1545 (2000)
11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
repeat_region <3..>108
/note="putative"
/rpt_family="MIR"
variation /rpt_type=dispersed
73
/frequency="0.01"
/replace="t"
repeat_region <207..>295
/note="putative"
/rpt_family="MIR"
variation /rpt_type=dispersed
308
/frequency="0.01"
/replace="t"
satellite <320..>339
/note="putative"
/rpt_type=tandem
repeat_region <340..>637
/note="putative"
/rpt_family="Alu"
variation /rpt_type=dispersed
471
/frequency="0.01"
/replace="g"
variation /frequency="0.01"
545
/replace="t"
variation /frequency="0.01"
560
/replace="t"
variation /frequency="0.22"
624
/replace="t"
variation /frequency="0.07"
/replace="c"
satellite <638..>718
/note="putative"
/rpt_type=tandem
variation 832
/frequency="0.45"
/replace="t"
mrna join(1060..1094,1855..1920,3013..3205,3786..4645)
/product="apolipoprotein-E"
variation 1163
/frequency="0.35"
/replace="c"
variation 1522
/replace="a"
variation /frequency="0.01"
1575
/frequency="0.01"
/replace="t"
CDS join(1878..1920,3013..3205,3786..4503)
/note="APOE"
/codon_start=1

/product="apolipoprotein-E"
/protein_id="AA627089.1"
/db_xref="GI:11034801"
/translation="MKVLAALLVPLACGAKVEQAVETEPPELRQOTEMOSGQRM
ELALGPMFLYLRWQTLSPNOEELISQVQETRALMETKELKAYSELEQLT
VAEETRLRSKELOAARLADMEVRCRLVQYREVOAMLGOSTEEELRLASLR
KPKRRLLRDADDQKRLAVYQAGARGAGRLSALREKGPLVEQGRVRAATVGSILAG
QPIQERAAQMGRLRAVMEQMSRTDRIDEVKEQVAEVRARLEQAOQIRLQAEAFQ
ARLKSWEPELVEDMQRMAGLVKEVQAAVGTSAAPVPSDNH"
1998
variation /frequency="0.10"
/replace="a"
repeat_region <2124..>2435
/note="putative"
/rpt_family="Alu"
variation /rpt_type=dispersed
2440
/frequency="0.21"
/replace="a"
repeat_region <2569..>2848
/note="putative"
/rpt_family="Alu"
variation /rpt_type=dispersed
2907
/frequency="0.01"
/replace="g"
variation /frequency="0.01"
3106
/replace="c"
repeat_region <3472..>3588
/note="putative"
/rpt_family="Alu"
variation /rpt_type=dispersed
3673
/frequency="0.01"
/replace="g"
variation /frequency="0.01"
3701-3702
/frequency="0.01"
/replace="ct"
3937
/frequency="0.14"
/replace="c"
variation /frequency="0.01"
4036
/replace="t"
variation /frequency="0.07"
4075
/replace="t"
repeat_region <4735..>5056
/note="putative"
/rpt_family="Alu"
variation /rpt_type=dispersed
5065..>5476
/note="putative"
/rpt_type=dispersed
5229-5230
/frequency="0.03"
/replace="gg"
variation 5229
/frequency="0.07"
/replace="t"
variation 5229-5230
/frequency="0.40"
/replace="g"
variation 5230
/frequency="0.13"
/replace="a"
variation 5361
/frequency="0.06"
/replace="c"
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

Alignment Scores:

Pred. No.: 4,32e-71
 Score: 1388.50
 Percent Similarity: 60.608
 Best Local Similarity: 60.608
 Query Match: 87.38%

Length: 5491
 Matches: 303
 Conservative: 0
 Mismatches: 3
 Indels: 194
 Gaps: 1

US-09-827-854-16 (1-317) x AF261279 (1-5491)

```

QY 12 PheLeuAlaGlyCysGlnAlaIysValGluGlnAlaValGluThrGluProGluProGlu 31
    |||
Db 3003 TTCACACAGAGATGCGAGGCCAAGGTGGAGCAAGCGGTGGAGACAGAGCCGAGCCGAG
QY 32 LeuArgGluGlnThrGluThrGluThrGluThrGluThrGluThrGluThrGluThrGluThr 51
    |||
Db 3063 CTGCGCCAGACAGCCAGGTGGAGAGAGCGCGCTGGAGACTGGACCTGCTGCTGCTTT
QY 52 TrpAspTyrLeuArgTyrPylGlnThrLeuSerGluGlnValGlnGlnGluLeuLeuSer 71
    |||
Db 3123 TGGGATTACCTGGCTGGGTGGAGACACTGTGAGCAGGTGGAGAGAGAGCTGCTCAGC
QY 72 SerGlnValThrGlnGluLeu----- 78
    |||
Db 3183 TTCCAGGTACCCAGAGACTGAGGTGAGTGCCTCCCTGGCCCTTGACCCCTCCTGCTG
QY 78 ----- 78
Db 3243 GCGGCTATACCTCCCGAGGTCCAGGTTTACCTTGCCTGCTGCTGCTGCTGCTGCTG
QY 78 ----- 78
Db 3303 CCTGGGTCTCTGCTGCTTCTAGCTTCTCTCCATTTCTGACTCCTGCTTCTAGCTTC
QY 78 ----- 78
Db 3363 TGGATTCTCTCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 3422
QY 78 ----- 78
Db 3423 CGTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3482
QY 78 ----- 78
Db 3483 CTCACGTGTGTGCGCCAGGCTGTGAACTCTGTGGCTCAAGCCATCCTCCCGCTCGG
QY 78 ----- 78
Db 3543 CCTCCCAAGTGTGGATTAGAGGATGAGCCACTTGCCTGCTCTAGCTCTCTCTCTCTCT 3602
QY 78 ----- 78
Db 3603 TCGTCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3662
QY 78 ----- 78
Db 3663 GCCTGCGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3722
QY 78 ----- 78
Db 3723 CCCCATCCAGCCCTTCTCCCGCTCCCATGCTGGACACCTCCCGCTCTCTGCGCGG 3782
QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 98
    |||
Db 3783 CAGGGGCTGATGAGAGACCATGAGAGGTTGAGAGGCTTACAAATCGAACTGTGAGGA 3842
QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAl 118
    |||
Db 3843 ACAACTGACCCCGGTGGCGGAGAGACGCGGCTGTCTCCAAAGAGCTGTGAGGCGGC 3902
QY 118 acGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgG 138
    |||

```

```

Db 3903 GCAGGCCCGCTGGGCGGAGACATGAGACGTGTGCGGCCGCTGTGCTCAGTACCGCGG 3962
QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerH 158
    |||
Db 3963 CGAGGTGACAGCCATCTGTGGACAGACACCGAGAGCTGTGGGTGCTGCTGCCCTCCA 4022
QY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAlaAspAlaAspAlaGlnLysCysLeuAl 178
    |||
Db 4023 CCTGCCAGAGCTGCTTACAGGCGCTCTCCGCATGCGCATGACCTGCACAAAGCGCTGGC 4082
QY 178 aValTyrGlnAlaGlyAlaArgGluGlyValGluGluArgGlyLeuSerAlaIleArgGluAr 198
    |||
Db 4083 AGTGTACAGCGCGGCGGCGGAGGCGCCGAGCGCGCTCAGCGCCATCCGCGCAGCGG 4142
QY 198 gLeuGlyProLeuValGluGlnGlyValArgValArgAlaIleThrValGlySerLeuAlaG 218
    |||
Db 4143 CTTGGGCGCCCTGTGTGGAACAGGCGCGTGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 4202
QY 218 yGlnProLeuGlnGluArgAlaGlnAlaIleThrGlyGluArgLeuArgAlaArgMetGluG 238
    |||
Db 4203 CCAGCGGCTACAGAGAGCGGCGCCAGGCTGCGGCGAGCGGCTGCGCGCGGATGAGAGA 4262
QY 238 uMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAr 258
    |||
Db 4263 GATGGGACAGCCGAGCCCGGACCGCTGTGAGCGAGGTGAAGAGACAGGTGGCGAGGTGGC 4322
QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaAr 278
    |||
Db 4323 CGCCAAAGCTGGAGAGACAGCCCAAGCATACGCTGCGAGCGCGAGCCCTTCCAGGCGCG 4382
QY 278 gLeuLysSerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalagLysLeuVa 298
    |||
Db 4383 CCTCAAGAGCTGTGTGAGCCCTGTGTGAAGACATGACAGCGCCAGTGGCGCGCTGTGT 4442
QY 298 lGluLysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||
Db 4443 GGAGAGAGTGCAGAGCTGCGTGGGACACAGCGCGCCCTGTGTGCCAGGACAAATCAC 4500

```

Search completed: March 14, 2003, 17:32:29
 Job time : 2314.23 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-16
Perfect score: 1589
Sequence: 1 MKVLMALLVTLFLAGCOAKV.....VEKVQAAVGSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPRO.spool/US09827854/runat_11032003.101609_27476/app_query.fasta.1.3576
-DB=Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NORR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.cgn.1.1.1201 @runat_11032003.101609_27476 -NCRP=6 -ICPU=3
-NO_XLPPY -NO_MMAB -LARGQUERY -NECL_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	22	AAF84314	Human ApoE2 coding
2	1589	100.0	1156	24	AAD22049	Human apolipoprotein
3	1582	99.6	1156	24	AAD22050	Human apolipoprotein
4	1577	99.2	1110	7	AAD26035	Human apolipoprotein
5	1577	99.2	1110	7	AAN60409	Human apolipoprotein
6	1577	99.2	1147	22	ABA83113	Human apolipoprotein E o
7	1577	99.2	1156	22	AAF84315	Human apolipoprotein
8	1577	99.0	1156	24	AAD22048	Human apolipoprotein
9	1573	99.0	1156	24	AAD22052	Human apolipoprotein
10	1569	98.7	1156	24	AAD22051	Human apolipoprotein
11	1568	98.7	1157	17	AAT06957	Human apolipoprotein
12	1568	98.7	1157	24	ABN95746	Human apolipoprotein
13	1568	98.7	1157	24	ABK64514	Gene #2244 used to
14	1568	98.7	1157	24	ABL65450	Human benign prost
15	1567	98.6	1110	6	AAN50450	Lung cancer relate
16	1565	98.5	1156	22	AAF84316	Sequence encoding
17	1565	98.5	1156	24	AAD22047	Human ApoE4 coding
18	1473.5	92.7	1279	22	AAS22437	Human apolipoprotein
19	1454.5	91.5	1107	19	AAK75756	Human cDNA encodin
20	1388.5	87.4	9360	24	ABL31915	Human ApoE genomic
21	1388.5	87.4	10716	24	AAD26034	Human apolipoprotein
22	1383	87.0	3805	20	AAZ08524	Human Apo E genomic
23	1383	87.0	3805	20	AAZ09526	Human Apo E genomic
24	1368.5	86.1	10716	24	AAD26108	Human apolipoprotein
25	1155.5	72.7	965	24	AAD32081	Bovine ApoE gene.
26	1155.5	72.7	5617	24	AAD32077	Human alpha-1-anti
27	1155.5	72.7	6026	24	AAD32075	Human alpha-1-anti
28	1130	71.1	1126	19	AAV29159	Human albumin prom
29	977	61.5	936	15	AAQ69101	Nucleotide sequenc
30	977	61.5	936	17	AAT18070	ApoE4Lx2 protease
31	976	61.4	660	18	AAT69792	Human ApoE4Lx2 cDN
32	939	59.1	597	17	AAT18068	Partial human apol
33	930	58.5	597	15	AAQ69096	Human ApoE4 cDNA.
34	876	55.1	1381	22	AA522673	ApoE4L protease cd
35	859.5	54.1	786	21	AAF18114	Human cDNA encodin
36	670	42.2	600	20	AAV89595	Lung cancer associ
37	651	41.0	407	24	ABK34238	EST clone CP147.
38	625	39.3	478	24	ABN95994	Human cDNA for nov
39	625	39.3	478	24	ABL62679	Gene #2492 used to
40	625	39.3	478	24	ABL67340	Colon adenocarcino
41	594	37.4	499	22	AA503045	Thyroid cancer rel
42	506	31.8	330	12	AAQ11980	Human diagnostic a
43	485	30.5	405	21	AAQ02139	Human apolipoprote
44			345	22	AAH98479	Human secreted pro
45	413	26.0	253	19	AAV39612	Human EST-derived
						E2 allele DNA enco

ALIGNMENTS

RESULT 1
AAF84314
AAF84314 standard; cDNA: 1156 BP.
AAF84314:
21-JUN-2001 (first entry)
Human ApoE2 coding sequence.
Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
Homo sapiens.
Key CDS
Location/Qualifiers
61..1014
/*tag= a
/product= "Human ApoE2"

XX JP2001017028-A.
 PN
 XX 23-JAN-2001.
 PD
 XX
 PF 28-APR-2000; 2000JP-0128919.
 XX
 PR 06-MAY-1999; 99JP-0125647.
 XX
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 DR WPI: 2001-285406/30.
 XX P-PSDB: AAB0996.
 DR
 XX New apoe humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 PS
 PS Disclosure; Page 11-12; 22pp; Japanese.
 CC
 CC The present invention relates to an APOE humanised mammalian cell. The
 CC present sequence is the coding sequence for human APOE2, which was used
 CC in the method of the present invention. The APOE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 CC
 XX Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

Alignment Scores:
 Pred. No.: 5 29e-116 Length: 1156
 Score: 1589.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-16 (1-317) x AAF84314 (1-1156)

QY 1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 DB 61 ATGAAGTTCTGTGGGCTGTGGTGTGTCACATTCCTGGCAGAGATGCCAGGCGAAGT 120
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGAGACAGAGCGGAGCCGAGCTGCGCAGAGACCGAGTGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaGlnLeuLeuSerGlnPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGGCAGCGGTGGAG 240
 QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAlaGlnAla 80
 DB 241 CTGTCTGACAGAGTGCAG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGACAGAGACCAATGAAGAGTGAAGGCTTACAAATCGGAACACTGAGAGAACTG 360
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerGlnLeuGlnAlaGlnAla 120
 DB 361 ACCCGCGGTGGAG 420
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 421 CGGCTGGGCGGAGACTGAG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCCAGAGACACGAGAGAGTGGGGTGGCTGCCCTGCCACCTGCGCC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValLys 180
 DB 541 AAGCTGCGTAACGGGCTCTCCGCGATGCCGATGACCTGCAGAAAGTGGCTGGAGAGTGTAC 600

QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGCGCGGGGCCCGCAGAGGCGCGCAGAGCGGCTCAGCGCCATCCGCGAGCGCCCTGGGG 660
 QY 201 ProLeuValGlnGlnGlnThrArgValArgAlaIleThrValGlySerLeuAlaGlnPro 220
 DB 661 CCCCTGTGGAAACAGGCGCGCGTGGGCGCCACTGTGTGGCTCCTGGCCGCGCAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGlyArgLeuArgAlaArgMetGluMetGly 240
 DB 721 CTACAGAGAGCGGGCCAGGCTGTGGCGAGCGGCTCCGCGCGATGGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB 781 AGCCGAGACCCGCGACCGCTGAGCAGAGTGAAGAGACAGAGTGGCGGAGGTGGCCCAAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGAGCGCCAGCAGATACCGCTGACAGGCGCCTTCCAGGCGCCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
 DB 901 AGCTGTGTGAGAGCCCTGTGTGGAAAGACATGACAGCGCAGTGGGCGGCTGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGAGCTGCGGTGGGACACAGCGCCGCCCTGTGCCAGGACACAAATCAC 1011

RESULT 2
 AAD22049
 ID AAD22049 standard; DNA; 1156 BP.

AC AAD22049;

DT 12-FEB-2002 (first entry)

XX Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.

XX Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;

KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.

OS Homo sapiens.

FT sig_peptide Location/Qualifiers

FT CDS 61..1014

FT /tag= a /product= "Human apoE isoprotein, apoE2"

FT mat_peptide 61..114

FT /tag= b

FT /tag= c /product= "Mature human apoE isoprotein, apoE2"

XX WO200177136-A1.

PN 18-OCT-2001.

PD 06-APR-2001; 2001WO-US11358.

PF 06-APR-2000; 2000US-0544386.

PR 04-OCT-2000; 2000US-0679088.

PR 05-APR-2001; 2001US-0827854.

XX (KOSP-) KOS PHARM INC.

PA (UYBO-) UNIV BOSTON.

XX Zannis VI, Kyriacos KE;

XX WPI: 2002-010885/01.
 DR P-PSDB: AAE13295.
 XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering

PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 XX
 PS Claim 14; Page 81-82; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoform, apoE2 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
 Alignment Scores:
 Pred. No.: 5,296-116 Length: 1156
 Score: 1589.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-16 (1-317) x AAD22049 (1-1156)
 QY 1 MetLeuValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 DB 61 ATGAAGCTTCTGGCGCTCGTTGCTGTCACATTCCTGGAGATGCCAGCCCAAGTG 120
 QY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnThrGlnSer 40
 DB 121 GAGCAAGCGGTGAGACACAGCGGAGCCGAGCTGCCAGACAGCCAGTGGCAGAC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCACACCTGGAGAACCTGCTGCTGCTTGTGGATTAACCTGGCGGTGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnLeuArgAla 80
 DB 241 CTGTCTGACAGAGTGGAGAGAGGAGCTGCTACCTCCAGTCAACCCAGGACTGAGGCGC 300
 QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 DB 301 CTGATGAGAGACACCATGAGAGAGTGAAGGCTTCAATTCGGAATGAGACACTG 360
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAla 120
 DB 361 ACCCCGCTGGCGAGAGACGCGGCGGCTGTCACAGAGCTGAGCGCGCGAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetLysPylLysGlyArgLeuValGlnTyrArgGlyVal 140
 DB 421 CGGCTGGCGGACATGAGAGAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCGATCTGCGCAGAGACACGAGGAGCTGCGGCTGCGCTGCCACCTGCGCC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 541 AAGCTCCGTAACCGGCTCTCCGCGATGCCGATGACCTCAAGAGGCTGCGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGlnGlnValArgGlyLysSerAlaHisLeuArgLeuGly 200
 DB 601 CAGGCGCGGCGCGCGAGGCGCGCGCGCTGAGCGCATCCGACGAGCGCTGCGG 660
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 661 CCCCTGTGTGAACAGGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240

DB 721 CTACAGAGACGGGCGCGCTGGGCGAGCGGCTGCCGCGCGATGGAGAGATGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaVal 260
 DB 781 AGCGGAGCGGCGCGCTGGAGCGAGTGAAGAGACAGTGGAGGAGGCGCGCGCGC 840
 QY 261 LeuGlnGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGAGAGACAGCGCGCAGAGATACCCCTGCAAGCGGAGCGCTTCCAGGCGCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 DB 901 ACTGCTTGCAGCCCTGCTGGAAGACATGACAGCGCCAGTGGCGCGCTGCTGGAAG 960
 QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGACAGCTGCGGTGGACACAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
 RESULT 3
 ID AAD22050 standard; DNA; 1156 BP.
 AC AAD22050;
 DT 12-FEB-2002 (first entry)
 DE Human apolipoprotein E (apoE) allele, apoE1 DNA.
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE1"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "mature human apoE allele, apoE1"
 FT WO200177136-A1.
 PN 18-OCT-2001.
 PD 06-APR-2001; 2001WO-US11358.
 PF 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 PI Zannis VI, Kyriacs KE;
 DR WPI: 2002-010885/01.
 DR P-PSDB; AAE13296.
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 PS Claim 14; Page 82; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE1 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)

CC for haplotyping and/or genotyping the APOE gene. The haplotyping
 CC methods of the invention are useful for improving the efficacy and
 CC reliability of several steps in the discovery and development of
 CC drugs for treating diseases associated with APOE activity, e.g.
 CC familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
 CC atherosclerosis, and Alzheimer's disease. They are useful to validate
 CC APOE as a candidate agent for treating a specific condition or disease
 CC predicted to be associated with APOE activity and in the design of
 CC clinical trials of candidate drugs for treating a specific condition
 CC or disease predicted to be associated with APOE activity. Genotyping
 CC or haplotyping methods are useful to screen for compounds targeting
 CC APOE to treat a specific condition or disease associated with APOE
 CC activity. The present sequence is a cDNA encoding human APOE protein.
 CC APOE gene is located on chromosome 19q13.2.

XX
 SQ Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;

Alignment Scores:

Pred. No.:	3,73e-115	Length:	954
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-16 (1-317) x AAN26035 (1-954)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 |||||
 Db 1 ATGAAAGCTTCTGGGCTCGTTCGTCGACATCTCTGGAGAGCCAGCCAAAGGTG 60
 OY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnProGlnSer 40
 |||||
 Db 61 GAGCAAGCGGTGGAGACAGACGCGGCGGAGCTGCCAGACAGCCAGTGGCAGACC 120
 OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpPaspTrpLeuArgTrpValGlnThr 60
 |||||
 Db 121 GGCACACGCTGGGAACCTGACCTGGCTTTGGATTAACCTGGCGGTGGGAGACA 180
 OY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||
 Db 181 CTGTCTGACAGAGTGGAGAGAGCTGCTCCACAGTCAACCCAGGAACTGAGGGCG 240
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 |||||
 Db 241 CTGATGAGACAGACCATGAGAGGTTGAGAGCTTACAAATCCGAACTGAGAGACACTG 300
 OY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 301 ACCCGGTGGCGAGAGACGCGGCGAGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 360
 OY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
 |||||
 Db 361 CGGCTGGGGCGGACATGAGAGCTGTGCGCGCTGTGCAGTACCGCGCGAGGTG 420
 OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 421 CAGGCCATCTCGGCCAGAGACCCAGAGGCTCGGCTGCTGCCCTCCACCTGGCC 480
 OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
 |||||
 Db 481 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACCTGCAGAAAGGCGCTGCGCAGTATC 540
 OY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 |||||
 Db 541 CAGGCGGGGGCCCGGAGGCGCCGAGCCGCTCAGCGCATCCGCGAGCGCTGGGG 600
 OY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||
 Db 601 CCCCTGTGTGAACAGGCGCGCTGCGCGCCGCTGCTCCCTGCGCGCCAGCGG 660
 OY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnGly 240
 |||||

Db 661 CTACAGAGCGGGCCAGCGCTGGGGCGAGCGGCTGCCGCGGATGAGAGATGGC 720
 OY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 |||||
 Db 721 AACCGGACCCGCGACCGCTGGAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 OY 261 LeuGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 |||||
 Db 781 CTGAGAGACAGCGCCAGCATACCGCTGCAGGCGGAGGCTTCCAGGCGCGCTCAAG 840
 OY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 |||||
 Db 841 ACCTGTTGAGACCCCTGCTGGAAGACATGACAGCGCCAGTGGCGGCTGTGTGAGAG 900
 OY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||
 Db 901 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTCGACGACAAATAC 951

RESULT 5

ID	AAN60409	standard: DNA; 1110 BP.
XX	AAN60409;	
AC	01-JAN-1980	(first entry)
DF		
DE	Human apolipoprotein-E.	
XX		
KW	Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.	
XX		
OS	Homo sapiens.	
XX		
EH	Key	Location/Qualifiers
FT	CDS	15..968
FT		/*tag= a

AU8547513-A.

24-APR-1986.

17-SEP-1985; 85AU-0047513.

PR 11-JUN-1985; 85JP-0126989.

XX (MITU) MITSUBISHI CHEM IND KK.

DR WPI: 1986-150217/24.

XX P-PSDB; AAP60507.

PT New DNA sequence coding for human apolipoprotein-E - and

XX expression vectors and transformed cells contg. it

XX Disclosure: Fig 2; 45pp; English.

CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.

SQ Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:

Pred. No.:	4.43e-115	Length:	1110
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	7	Gaps:	0

US-09-827-854-16 (1-317) x AAN60409 (1-1110)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20

DB 15 ATGAAAGTTCTGTGGCTGTGGCTGTGCACATTCCTGGCAGAGATCCAGGCGCAAGGCTG 74
 QY 21 GUGGAlAValAGlunthrgluprogiluprogilunleuArgInglunthrgluprglinsr 40
 DB 75 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGCCAGCAGACCGAGTGGCGAGC 134
 QY 41 gylglnArgrTpplunleuAlaLeuglArgrPheTrpAspTyrleuArgrTpValglnthr 60
 DB 135 GGGCAGCGGTGGGAACTGGGACTGGGTGGCTTTGGGATTACTGGCGGTGGTGCAGACA 194
 QY 61 LeuSerGlnGlnValGlnGlnGlnleuLeuSerSerGlnValThrGlnGlnleuAlaGAla 80
 DB 195 CTGTCTGAGACAGGTGAGAGAGAGTGTCTCAGCTCCAGGTCCACGAGAACTGAGAGCGC 254
 QY 81 LeuMetAspqlunthrmetylsGlnleuLysAlaTyrlySerGlnleuGlnGlnGlnleu 100
 DB 255 CTGATGAGACAGACCATGAAGAGGTGAAAGCCTACAAATCGGAACCTGAGAGAACAACTG 314
 QY 101 ThrProValAlaGlnGlnGlnthrglArgrAlaArgLeuSerLysGlnleuGlnAlaAlaGlnAla 120
 DB 315 ACCCTGGTGGCGGAGAGACGCGGCGACGCTGTCCAAGAGCTGCAGAGCGCGCGAGGCC 374
 QY 121 ArgLeuGlnAlaAspMetGlnAspValLysGlnArgLeuValGlnThrArgGlnGlnVal 140
 DB 375 CGGCTGGGCGCGACATGAGAGAGTGTGCGCGCTGTGTGAGTACCGCGGCGAGGCTG 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerThrLysLeuArg 160
 DB 435 CAGGCGCATGCTGGCGCCAGAGACACGAGAGACCTGGGCTGGCTCCCTCCACCTGCGC 494
 QY 161 LysLeuArgrLysArgrLeuLeuArgrAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 DB 495 AAGCTCGTAAAGCGGCTCCCTCCGATGCGCATGACCTCGAAGCGCGTGCAGAGTAC 554
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 DB 555 CAGGCGGCGGCGCGCGAGGCGCGAGCGCGCTCAGCGCCCTCCGCGAGCGCGCTGGGG 614
 QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 DB 615 CCCTGGTGGAAACAGGCGCGCGCTGGCGCGCGCGCTGGCGCTCCCTGGCGCGCGCGC 674
 QY 221 LeuGlnGlnAlaArgAlaGlnAlaTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 240
 DB 675 CTACAGAGAGCGGCGCGAGGCTGGGCGAGCGCGCTGGCGCGCGGTGGAGAGTGGGG 734
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnGlnGlnGlnGlnGlnGlnGln 260
 DB 735 AGCGGAGACCGCGACCGCGCTGGCGAGGTGAAGAGACAGGTGGCGAGGTGGCGCGC 794
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 DB 795 CTGAGAGAGAGCGCGCGAGCATGACCTGACGCGCGCGCGCTCCACAGGCGCGCGCTCAAG 854
 QY 281 SerTrpPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 300
 DB 855 AGCTGTGTGAGAGCGCGTGGTGAAGACATGACCGCGAGTGGCGCGCGGTGGTGAAG 914
 QY 301 ValGlnAlaAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 317
 DB 915 GTGCAAGCTGCGCTGGGCGACAGCGCGCGCTGTGCGCCAGCAGCAATCAC 965

RESULT 6
 ABA83113
 ID ABA83113 standard; DNA; 1147 BP.

AC ABA83113;
 DT 08-FEB-2002 (first entry)
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.

KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200175177-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 03-APR-2001; 2001WO-US10947.
 XX
 PR 03-APR-2000; 2000US-194336P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Morin RJ, Sherman-Baust CA, Pizer ES, Hough CD;
 XX WPI; 2001-626450/72.
 DR P-PSDB; ABB50287.
 XX
 PT Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene .
 XX
 PS Claim 23; Page 105-106; 140pp; English.
 XX
 CC The invention relates to methods for diagnosing and prognosing ovarian
 CC tumors in an individual via the detection and measurement of the
 CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer. In prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer and in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour. The ovarian tumour marker genes of the invention were identified
 CC using SAGE (serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 CC the invention.
 CC
 SQ Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;
 XX
 XX
 Alignment Scores:
 Pred. No.: 4.6e-115 Length: 1147
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 22 Gaps: 0

```

US-09-827-854-16 (1-317) x AAF84315 (1-1147)
OY      1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
        |||
DB      46 ATGAAGGTTCTGTGGGCTGGCTGGTGTGTCACATTCTCTGGCAGGATGCCAGGCCAAAGGTG 105
OY      21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
        |||
DB      106 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCGACAGACAGACGAGTGGCAGAGC 165
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
        |||
DB      166 GGCACGCGCTGGGAACCTGGACCTGGTGGCTTTGGATTACCTGCGTGGGTGGCAGACA 225
OY      61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
        |||
DB      226 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGATCACCAGAACTGAGGGCG 285
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
        |||
DB      286 CTGATGACAGAGACCATGAGAGATTGAAGGCCCTCAAAATCGGAAGCTGGAGGACAACTG 345
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
        |||
DB      346 ACCCGGTGGGAGAGACCGGGGACGCGCTGTCAAAGAGACTCAGCGCGCGCAGAGCC 405
OY      121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
        |||
DB      406 CGCGTGGGCGCGACATGAGAGAGCTGTGCGGCCCTGTGTGCTACCTCCGCGCAGAGTG 465
OY      141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
        |||
DB      466 CAGGCGATGCTCGGCGCAGAGACCGAGAGGCTGGGGTGGCGCTGCCCTCCACTGGGCG 525
OY      466 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
        |||
DB      526 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGCATGACGCAAGAACGCGCTGCGCACTGTAC 585
OY      181 GlnAlaGlyAlaArgGlnGlnValGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
        |||
DB      586 CAGCGCGGGGCGCGGAGGAGGCGCGGCTGACGCGCCATCCGAGCGGCGCTGGGGG 645
OY      201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
        |||
DB      646 CCCCTGGTGGAAACAGAGCGCGGTGGCGCGCACTGTGGGTCTGCGCGCGCAGAGCG 705
OY      221 LeuGlnGluArgAlaGlnAlaTrpGlyLeuArgLeuArgAlaArgMetGluGluMetGly 240
        |||
DB      706 CTACAGGAGCGGCGCCAGGCGCTGGGCGCAGCGGCTGCGCGCGGATGAGAGAGATGGGC 765
OY      241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
        |||
DB      766 AGCCGAGACCGCGACCGCTGGACAGGTGAAGAGAGAGGTGGCGGAGTGGCGCCAAAG 825
OY      261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
        |||
DB      826 CTGGAGGAGCAGCGCCACAGATAGCTCGAGCGCGGAGGCGCTTCAAGCCGCTCAAG 885
OY      281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
        |||
DB      886 AGCTGTGTCGAGCCCTGCTGGTGAACATGACAGCCGCACTGGGCGCGCTGTGTGAGAGA 945
OY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
        |||
DB      946 GTGCAGGCTGCGGTGGGACCAACGCGCCCTGTGTGCCAGGAGCAATAC 996

```

```

DT      21-JUN-2001 (first entry)
XX
DE      Human ApoE3 coding sequence.
XX
KW      Human; ApoE3; Alzheimer's disease; arteriosclerosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      CDS 61..1014
FT          /tag="a
          /product="Human ApoE3"
XX
PN      JP2001017028-A.
XX
PD      23-JAN-2001.
XX
PR      28-APR-2000; 2000JP-0128919.
XX
PR      06-MAY-1999; 99JP-0125647.
XX
PA      (MITU ) MITSUBISHI CHEM CORP.
XX
DR      WPI: 2001-285406/30.
XX      P-PSDB: AAB80997.
XX
PT      New apoE humanized mammalian cell useful for screening for agents
PR      useful for treating or preventing Alzheimer's disease and
PT      arteriosclerosis -
XX
PS      Disclosure: Page 13-14; 22pp; Japanese.
XX
CC      The present invention relates to an ApoE humanised mammalian cell. The
CC      present sequence is the coding sequence for human ApoE3, which was used
CC      in the method of the present invention. The ApoE humanised mammalian cell
CC      can be used for screening for agents useful for treating or preventing
CC      Alzheimer's disease and arteriosclerosis.
XX
SQ      Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other:
SO

```

Alignment Scores:

Pred. No.:	Length:	Matches:	Score:
1577.00	1156	316	99.68%
Best Local Similarity:		99.68%	
Query Match:		99.24%	
DB:	22	Gaps:	0

```

US-09-827-854-16 (1-317) x AAF84315 (1-1156)
OY      1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
        |||
DB      61 ATGAAGGTTCTGTGGGCTGGCTGGTGTGTCACATTCTCTGGCAGGATGCCAGGCCAAAGGTG 120
OY      21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
        |||
DB      121 GAGCAAGCGGTGAGACAGAGCCGAGCTGCCGACAGACAGACGAGTGGCAGAGC 180
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
        |||
DB      181 GGCACGCGCTGGGAACCTGGACCTGGTGGCTTTGGATTACCTGCGTGGGTGGCAGACA 240
OY      61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
        |||
DB      241 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGATCACCAGAACTGAGGGCG 300
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
        |||
DB      301 CTGATGACAGAGACCATGAGAGATTGAAGGCCCTCAAAATCGGAAGCTGGAGGACAACTG 360
OY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
        |||
DB      361 ACCCGGTGGGAGAGACCGGGGACGCGCTGTCTCAAGAGAGCTGCAAGCGCGCGCAGAGCC 420

```

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 DB 421 CGGCTGGGCGGACATGAGACGTGTGGGGCCGCTGTGACGATACCGCGGAGG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 481 CAGGCCATGCTCGGCGACAGACCGAGAGCTGGGGTGGCTCCCTCCACCTGGCCG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||||
 DB 541 AACCTCGTAAGCGGCTCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 DB 601 CAGGCGGGGCGCGAGGCGCGGAGCGGGCTGCAGCGCCATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 DB 661 CCCCTGGTGGAAACAGGCGCGCTGCGGGCCGACCTGTGGGCTCCTGGCGGCGCAGCG 720
 QY 221 LeuGlnGluArgGluAlaArgTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 |||||||
 DB 721 CTACAGAGAGCGGCGCCAGGCTGTGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||||
 DB 781 AGCGGACCGCGGACCGCTGCGAGAGTGAAAGAGACAGGTGGGAGGTGGCGGCAAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||||||
 DB 841 CTGGAGAGACAGGCGCCAGACGATACCTGCGAGCGCGCTTCCAGGCGCCGCTCAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||||
 DB 901 AGCTGGTTGAGAGCCCTGTGGGAAGACATGCAGCGCCAGTGGGGCGGCTGTGGAGAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||||
 DB 961 GTGCAGAGCTGCGGTGGGACACAGCGCGCGCTGTGCCAGCAGACATCAG 1011
 RESULT 8
 AAD22048
 ID AAD22048 standard; DNA; 1156 BP.
 AC AAD22048;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 XX
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE isoprotein, apoE3"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 XX
 PN WO200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR

PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypreos KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB: AAE13294.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridemia
 XX
 PS Claim 14; Page 81; 91pp; English.
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
 Alignment Scores:
 Pred. No.: 4 64e-115 Length: 1156
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-16 (1-317) x AAD22048 (1-1156)
 QY 1 MetLysValLeuTrpAlaIleLeuValIlePheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 DB 61 ATGAAGTTCTGTGGGCTGCTGCTGTCACATTCCTGGCAGAGATGCCAGGCGAAG 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnIleThrGluTrpGlnSer 40
 |||||||
 DB 121 GAGCAAGCGGTGGAGACAGACGCGGAGCCGAGCTGCGCAGACAGACCAAGTGGCAGAC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGluLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 DB 181 GGCAGAGCTGGGAACTGGCACTGGGTCCCTTTTGGGATTACCTGGCTGGGTGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValIleThrGlnGluLeuArgAla 80
 |||||||
 DB 241 CTGTCTGACAGAGTGCAGAGAGAGCTGTCACTGCCAGGTCAACCAGAACTGAGAGCGG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
 |||||||
 DB 301 CTGATGACGAGACCAATGAAGAGTGAAGGCTTACAAATCGAATCGAGAACAACTG 360
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 DB 361 ACCCGGTGGCGAGAGAGCGGGGCGAGCTGTCCAAAGAGCTGCAGGCGGCGCGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 DB 421 CGGCTGGGCGGACATGAGAGACGTGTGGGGCCGCTGTGACGATACCGCGGAGG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 481 CAGGCCATGCTCGGCGACAGACCGAGAGCTGGGGTGGCTCCCTCCACCTGGCCG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||||

Db 541 AAGCTGCGTAAGCGCTCCTCCGCGATGGCGATGACCTGCAGAGCGCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
Db 601 CAGGCGGG 660
QY 201 ProLeuValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 220
Db 661 CCCCTGTGTAACAGAGCG 720
QY 221 LeuGlu 240
Db 721 CTACAGGAGCGGG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGluGluGluGluGluGluGluGluGlu 260
Db 781 ACCCGACCG 840
QY 261 LeuGlu 280
Db 841 CTGAGAGGAGCAGCG 900
QY 281 SerTTPheGluProLeuValGluAspMetGluArgGluGluGluGluGluGluGluGluGlu 300
Db 901 AGCTGTTCGAGCG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAAGGCTGCGGTGGGACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1011
RESULT 9
AAD22052
ID AAD22052 standard; DNA: 1156 BP.
AC AAD22052;
XX 12-FEB-2002 (first entry)
DT
XX Human apolipoprotein E (apoE) allele, apoE2** DNA.
DE
XX Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 61..1014
FT /*tag= a
FT /product= "Human apoE allele, apoE2**"
FT sig_peptide 61..114
FT /*tag= b
FT mat_peptide 115..1011
FT /*tag= c
FT /product= "Mature human apoE allele, apoE2**"
PN WO200177136-A1.
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-US11358.
PF
XX
XX 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
XX Zannis VI, Kyriacos KE;
PI WPL: 2002-010885/01.
DR P-PSDB: AAE13298.
XX

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
PS Claim 14; Page 83; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2** DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;
Alignment Scores:
Pred. No.: 9,56e-115 Length: 1156
Score: 1573.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 98.99% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-16 (1-317) x AAD22052 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGGCTCGTTGCTGTCACATTCCTGCAGAGATCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValAlaGluThrLeuProGluProGluLeuArgGluGluThrLeuTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnTrp 60
Db 181 GGCACAGCGCTGGGAACCTGGCACTGGGTGCTTTGGGATTACCTGGCGGGGCGAGACA 240
QY 61 LeuSerGluGlnValGluGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTGACAGAGTGCAGAGAGAGCGTGCACCTCCAGTCACCCAGGAACTGAGAGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
Db 301 CTGATGAGACGAGACCATGAAGAGTTGAAGGCTTCAAAATCGAAGTGAAGACAACATG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGAGCG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuLaserHisLeuArg 160
Db 481 CAGGCCATGCTCGCGCAGACACCGAGCGAGCTGGGGTGGCTGCCTCCACCTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 541 AAGCTGCGTACAGCGGCTCCTCGCATGGCATGACCTCAGAAAGCGCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
Db 601 CAGCGCGGG 660
QY 201 ProLeuValGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 220
Db 661 CCCCTGTGTAACAGAGCG 720
QY 221 LeuGlu 240

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpPalaGlyLeuValGluLys 300
 |||||
 Db 901 AGTGGTTGAGCCCTGGTGGAGACATGACAGCCAGTGGGCGCGCTGGTGGAGAG 960
 QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||
 Db 961 GTGCAGGCTGCGCTGGGACACAGCCGCCCTGTGCTCCAGCAGCAATAC 1011
 RESULT 11
 AAT06957
 ID AAT06957 standard: cDNA to mRNA: 1157 BP.
 AC AAT06957;
 XX
 DT 19-JUN-1996 (first entry)
 XX
 DE Human apolipoprotein-E (APOE) cDNA.
 XX
 DE Recombinant; human; apolipoprotein-E; APOE; insect cells; larva;
 KW Manduca sexta; Autographica californica nuclear polyhedrosis virus;
 KW haemolymph; lipid complex; biologically active; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc_feature 62..119
 FT note="misc_signal"
 FT 62..1015
 FT /*tag= b
 FT CDS
 XX US472858-A.
 XX
 PN 05-DEC-1995.
 XX
 PD 04-JUN-1991: 91US-0709949.
 XX
 PF 04-JUN-1991: 91US-0709949.
 XX
 PR 04-JUN-1991: 91US-0709949.
 XX
 PA (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Attie AD, Beckage NE, Gretsch DG, Sturley SL;
 XX
 DR WPI: 1996-029812/03.
 DR P-PSDB: AAR86791.
 XX
 PT Prodn. of recombinant apo:lipoprotein E in insects - by infecting
 PT Manduca sexta larvae with recombinant Autographica californica
 PT nuclear polyhedrosis baculovirus vector.
 XX
 PS Disclosure: Columns 11-14: 10pp; English.
 XX
 CC Recombinant human apolipoprotein-E (APOE) (AAR86791) can be produced
 CC by preparing a genetic construct (contg. an APOE-encoding sequence,
 CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
 CC to be expressed in insect cells), which is then introduced into a
 CC Manduca sexta larva (using a recombinant Autographica californica
 CC nuclear polyhedrosis virus) and recovering the protein from the
 CC haemolymph of the larval host. The APOE produced is in a form
 CC sufficiently complexed with lipids to be biologically active, which
 CC cannot be achieved in insect cell cultures, and can therefore be
 CC used in therapeutic applications.
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 2,366-114
 Score: 1568.00
 Percent Similarity: 99.05%
 Best Local Similarity: 99.05%
 Query Match: 98.68%
 DB: 17
 Length: 1157
 Matches: 314
 Conservative: 3
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-827-854-16 (1-317) x AAT06957 (1-1157)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 62 ATGAAGGTTCTGGGCTCGTGGTGGTGCATCTTCTGGCAGAGGCCAGGCAAGGTG 121
 QY 21 GlnGlnAlaValAlaGlnTrpGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
 |||||
 Db 122 GAGCAAGCCGTGGAGACAGACAGCCGAGGCCGAGCTCCGACAGACCCAGTGGCAGAGC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 Db 182 GGCACAGCCTGGGAACCTGGACACTGGGTGCTTGTGGATTACTGCGGTGGGTGAGACA 241
 QY 61 LeuSerGlnGlnValGlnGlnGluLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||
 Db 242 CTGTCTGACAGAGTGCAGAGAGGAGCTGCTCAGCTCCCAAGTCAACCAAGAACTGAGAGGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
 |||||
 Db 302 CTGATGAGACAGACCATGAAGAGTTGAAGCCTACAAATCGAAGCTGAGAGAACACTG 361
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 362 ACCCCGCTAGCGGAGAGAGACGCGGCGCTGTCCAAAGAGCTGCAGACGCGCAGGCC 421
 QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
 |||||
 Db 422 CGGCTGGGCGGACATGAGAGAGCTGCTGCGGCGCTGTGATGATACCGCGGCGAGGTG 481
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 482 CAGCCATGCTCGGCGCAGAGACACCGAGAGACTCGGGTGGCTCCCTCCCACTGGCCG 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||
 Db 542 AAGCTCGTAAGGCGCTCTCCGCGATCCGATGATGATGATGATGATGATGATGATGATGATGAT 601
 QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
 |||||
 Db 602 CAGGCGGGGCGCGGAGGCGCGGCGCGGCGGCTGAGCGCATCCGCGCAGCGCCGCGGGG 661
 QY 201 ProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||
 Db 662 CCCCTGTGTGAACAGGCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 |||||
 Db 722 CTACAGAGACGGGCGCGGCTGGGGCGAGCGGCTGCGCGCGGCGGATGAGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 |||||
 Db 782 AGTGGACCCGCGACCGCTGGAGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 |||||
 Db 842 CTGGAGAGACAGGCGCGAGATAGCCCTGACAGCGCGAGAGCGCTCCAGGCGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpPalaGlyLeuValGluLys 300
 |||||
 Db 902 ACTGTGTTGACGCCCTGTGTGAAGACATGACAGCGCCCAATGGGCGCGCTGTGTGGAGAG 961
 QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||
 Db 962 GTGCAGGCTGCGCTGGGACACAGCCGCCCTGTGCTCCAGCAGCAATAC 1012
 RESULT 12
 ABN95746
 ID ABN95746 standard: DNA: 1157 BP.
 XX
 AC ABN95746;
 XX
 DT 13-AUG-2002 (first entry)


```

XX  Sequence encoding human apolipoprotein E.
DE  Hyperlipaemia; ds.
XX
XX  Homo sapiens.
OS
XX
XX  Key      Location/Qualifiers
FH  CDS      15..968
FT          /*tag= a
FT          69..965
FT          mat_peptide
FT          /*tag= b
XX
XX  JP60118189-A.
XX
XX  25-JUN-1985.
XX
XX  29-NOV-1983; 83JP-0224980.
XX
XX  29-NOV-1983; 83JP-0224980.
XX
XX  29-NOV-1983; 83JP-0224980.
XX
XX  (MITU ) MITSUBISHI CHEM IND KK.
XX
XX  WPI: 1985-188003/31.
XX  P-PSDB: AAP51204.
XX
XX  DNA fragment - contg. DNA which codes human apolipoprotein E for
PT  treatment of hyperlipaemia.
XX
XX  Claim 3; Page 484; 8pp; Japanese.
XX
XX  The sequence may be used to produce the apolipoprotein E, useful in the
CC  treatment of hyperlipaemia.
XX
SQ  Sequence 1110 BP, 198 A; 355 C; 414 G; 143 T; 0 other;

Alignment Scores:
Pred. No.:      2,7e-114
Score:          1567.00
Percent Similarity: 99.05%
Best Local Similarity: 99.05%
Query Match:    98.62%
DB:             6
                Length: 1110
                Matches: 314
                Conservative: 0
                Mismatches: 3
                Indels: 0
                Gaps: 0

US-09-827-854-16 (1-317) x AAN50450 (1-1110)
OY  1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIAspVal 20
DB  15 ATGAAGGTTCTGTGGGCTCGTTGCTGGTCACATTCCTGCGAGATGCCAGGCCAAGTG 74
OY  21 GluGlnAlaIValGluThrGluProGluProGluLeuAlaGlnGlnThrGluTrpGlnSer 40
DB  75 GAGCAAGCGGTGGAGACAGACGCGGAGCCGAGCTGCCGACAGACCCAGAGCCAGAGAC 134
OY  41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB  135 GGCACAGCGGTGGAACTGGCACTGGCTCTTTGGGATTACCTGCCCTGGGTGCAGACA 194
OY  61 LeuSerGluGlnIValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB  195 CTGCTGACACAGCTGCAGAGAGCTGCTCACGCTCCAGAGTCAACCAGAACTGAGGGCG 254
OY  81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
DB  255 CTGATGGAGACCATGAGAGGTTGAAGGCTTACAAATCGAACTGAGAGAACACACTG 314
OY  101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIAspVal 120
DB  315 ACCCGGTGGCGGAGAGACGCGGCGCTGCTCCAGAGAGCTGCAGCGGCCAGGCC 374
OY  121 ArgLeuGlyAlaAspMetCysIAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
DB  375 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCCCTGTCAGTACCGCGCGAGGTG 434

```

```

OY  141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB  435 CAGGCCATCTGGGCCAGACACCGAGAGCTGGGGGTGCTGCCCTGCCACCTGCCG 494
OY  161 LysLeuArgLysArgLeuLeuArgAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB  495 AAGCTGCGTAAGCGGCTCTCCGCGATGACCTGCAGAAAGGCGCTGGCAGGTAC 554
OY  181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB  555 CAGGCGGGGCGCGGAGGGCGCCGAGCCGCGGCTCAGGCCATCCGAGAGCGCTGGGG 614
OY  201 ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyGlnPro 220
DB  615 CCCCTGTGGTAACAGGCGCGCTGGGGCGGCGGCTGTGGCTCCCTGGCGGCGGCGG 674
OY  221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB  675 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGGATGAGAGAGATGGGC 734
OY  241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB  735 AGCCGAGCCCGCGACCGCTGGACGAGGTGAAGAGAGCAAGTGGCGAGGTGGCGCCAG 794
OY  261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB  795 CTGGAGAGACCGAGCCAGATACGCTGCAGGCGCGAGGCTTCCAGGCCCGCTCAAG 854
OY  281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB  855 ACTGCTGTGACCCCTGCTGGAGACATGACAGCGCCGATGGGCGGCTGGTGGAGAAG 914
OY  301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB  915 GTCAAGCTGCGGTGGGACACAGCGCCGCTGTGCCAGACATATCAC 965

```

Search completed: March 14, 2003, 12:43:10
Job time : 183.728 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589
Sequence: 1 MKVLMAALLVFLAGCQAKV.....VEKQAAVGTSAAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xl
-Q=/cgn2_1/USPTO_spool/US09827854/rnat_11032003_101611_27510/app_query.fasta.1.3576
-DB=Issued_Patents.NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 -CGN1_1_1224 -rnat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents.NA:*
2: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/BACKFILES1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1568	98.7	1157	1	US-07-709-949-1
2	1130	71.1	1126	4	US-08-949-155-5
3	1130	71.1	1126	4	US-09-819-964-5
4	994	62.6	4267	4	US-08-949-155-51
5	994	62.6	4267	4	US-09-819-964-51
6	976	61.4	660	2	US-08-726-306A-28
7	503	31.7	330	1	US-07-849-389-6
8	374	23.5	252	4	US-08-617-256-24
9	374	23.5	252	4	US-09-287-141-24
10	374	23.5	252	4	US-09-431-613-24
11	374	23.5	252	4	US-09-504-245-24
12	374	23.5	252	4	US-09-287-682-24

13	374	23.5	252	4	US-09-287-679-24
14	374	23.5	252	4	US-09-397-766-24
15	374	23.5	252	4	US-09-287-681-24
16	374	23.5	252	4	US-09-495-444-24
17	176.5	11.1	842	1	US-08-952-796-1
18	173.5	10.9	801	1	US-07-959-946-4
19	173.5	10.9	801	1	US-08-333-577-4
20	173.5	10.9	801	5	PCT-US92-08634-4
21	159	10.0	5661	4	US-08-938-105-2
22	156.5	9.8	964	1	US-08-448-606-5
23	153.5	9.7	863	1	US-08-448-606-7
24	152.5	9.6	4296	4	US-09-060-410-3
25	150	9.4	3256	2	US-08-968-751-3
26	149	9.4	1879	4	US-09-750-580-2
27	146.5	9.2	603	4	US-08-952-796-14
28	146.5	9.2	4852	1	US-07-853-913-3
29	144	9.1	13121	4	US-08-961-527-16
30	141	8.9	6306	1	US-08-466-390-3
31	141	8.9	6306	1	US-08-470-950-3
32	141	8.9	6306	1	US-08-467-781-3
33	141	8.9	6306	1	US-08-195-487-3
34	141	8.9	6306	2	US-08-483-924-3
35	141	8.9	6306	5	PCT-US93-06160-3
36	141	8.9	8789	1	US-08-328-254-5
37	141	8.9	10136	5	PCT-US95-16216-2
38	141	8.9	10136	5	US-08-353-700-2
39	140	8.8	1771	2	US-08-533-669A-7
40	140	8.8	1771	2	US-08-511-872-1
41	140	8.8	1771	4	US-09-183-861-7
42	140	8.8	1771	4	US-09-022-765-7
43	140	8.8	1960	2	US-08-533-306A-1
44	140	8.8	1960	2	US-08-742-923A-1
45	140	8.8	4868	1	US-08-139-937-12

ALIGNMENTS

Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 126, App
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 12, Appl

RESULT 1
US-07-709-949-1
; Sequence 1, Application US/07709949
; Patent No. 5472858
; GENERAL INFORMATION:
; APPLICANT: Attie, Alan D
; APPLICANT: Gretch, Daniel G
; APPLICANT: Sturley, Stephen L
; APPLICANT: Beckage, Nancy E
; TITLE OF INVENTION: Production of Recombinant Proteins in
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/709,949
; FILING DATE: 19910604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 1:


```

; NAME/KEY: CDS
; LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 2,18e-101 Length: 1126
Score: 1130.00 Matches: 223
Percent Similarity: 83.44% Conservative: 44
Best Local Similarity: 69.69% Mismatches: 45
Query Match: 71.11% Indels: 8
DB: Gaps: 3

US-09-827-854-16 (1-317) x US-08-949-155-5 (1-1126)
OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGTCTTGTGGTAACTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCGCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158
OY 38 TrpGlnSerGlyGlnAlaTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrp 57
DB 159 TGGCAGGCGCAGCAGCCCTGGGAGCAGCCCTGCGGCTCTGCGATTACCTGCGCTGG 218
OY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 219 GTGCGTCCCTGTCTGACCAAGTCAGAGAGAGTGTCTACACCAAGTCACCCAGGAA 278
OY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlu 97
DB 279 CTGACGAGAGCTGATGAGAGAGAGCATGAAGAGGTGAAGCCCTACCGGAGAGAGCTGAG 338
OY 98 GlnGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
DB 339 GCGCAGCTGGGCGCCGTCACCCAGCAGCAGCGCGCCCTCTCCAAAGAGCTGACAGCG 398
OY 118 AlaGlnAlaArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArg 137
DB 399 GCGCAGGCGCCGCTGGGCGGCGACATGAGAGCGTGCACACCGCTTGTGCTCTACCGC 458
OY 138 GlyLysValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSer 157
DB 459 AGCGAGGTGCACAACATGTTGGGCGCAGACACCGAGAGCGTGGAGCGCCCTGCTTCC 518
OY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
DB 519 CACCTGCGCAGCTGCGCAGCGGCTGCTCCGCGACACCGAGAGAGCTGACAGAGCGCTG 578
OY 178 AlaValTrpGlnAlaGlnAlaArgGlnGlyArgGlyLeuSerPheAlaIleArgGlu 197
DB 579 GCGGTGTACCAAGCGGCGGCTGCGCAGGCGCGCAGCGAGTGTAGGCTCCGCGCAG 638
OY 198 ArgLeuGlyProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAla 217
DB 639 CGCCTCGGCGCCCTGTGTGTGAGAGCGCGCATTTGGCGCGCCACCTAGTACACGAGGC 698
OY 218 GlyLysProLeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGlu 237
DB 699 GCGCAGCGCTGCGGAGCGCGCGGCAAGCTGCGGCGCAGAAAGCTGCGGCGCTGAG 758
OY 238 GluMetGlySerArgTrpArgAspArgLeuAspGluValLysGlnGluValAlaGlnVal 257
DB 759 GAGATGGGCGACCGGACCGCGGACCGCTGATGATGATGATGATGATGATGATGATG 818
OY 258 ArgAlaLysLeuGluGlnGlnAlaGlnGlnLysArgLeuGlnGlnAlaAlaPheGlnAla 277
DB 819 CGCACAAGGTGAGAGAGCAGGCGCAGATTGGCGCTGCGAGCGCGAGAGATTCCAGCC 878
OY 278 ArgLeuLysSerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
DB 879 CTCCTCAAGAGCTGTGTCAGCTCTCTGTGTGAGACATATACGCGCAGTGGCGGCGCTG 938

```

```

OY 298 ValGlnLysValGlnAlaAlaValGly---ThrSerAlaAlaProValProSerAspAsn 316
DB 939 GTGAGAGAGATGACGTGCGCGGTGAGATATAGCTCTCCACCTTGCGCCCAAGTATAT 998

RESULT 3
US-09-819-964-5
; Sequence 5, Application US/09819964
; Patent No. 6369294
; GENERAL INFORMATION:
; APPLICANT: Piedrahita, Jorge A
; Bazer, Fuller W
; TITLE OF INVENTION: Compositions and Methods for the
; Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/819,964
; FILING DATE: 28-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/949,155
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1126 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51..1001
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:
Pred. No.: 2,18e-101 Length: 1126
Score: 1130.00 Matches: 223
Percent Similarity: 83.44% Conservative: 44
Best Local Similarity: 69.69% Mismatches: 45
Query Match: 71.11% Indels: 8
DB: Gaps: 3

US-09-827-854-16 (1-317) x US-09-819-964-5 (1-1126)
OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 51 ATGAGGGTTCTGTGGTCTTGTGGTAACTCTCTCGCAGATGCGGACAGAGAC 110
OY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
DB 111 GAGCCGGGG-----CCGCGCGGAGGTGCACGTGTGTGGAGAGAGCCCAAG 158

```

```

QY 38 TTPGInserGlyInArtrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTTP 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGCAGCCAGCCCTGGGAGCAGCCCTGGCCCTTCTGGGATTACCGCCCTGG 218
QY 58 ValGInThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTCCAGTCCCTGTCTACCAAGTCAGAGAGAGCTGCTCAGCCACCAAGTCCACCCAGGA 278
QY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATAGAGAGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 338
QY 98 GlnGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGCAGGAGCCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398
QY 118 AlaGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGCAGGAGCCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 458
QY 138 GlyGlnValGlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCGAGAGTGCACAAATGTTGGGCCAGACACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 518
QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTCGCAGAGCTGCGCAAGCGCGCTGCTCCGCGACACCGAGAGAGAGAGAGAGAGAG 578
QY 178 AlaValTyrGlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCGCTTACAGAGCGGAGCGCTGCGCGAGGCGCGAGCGCGAGCGCGAGCGCGCGCGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCGCTGGGAGCGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
QY 218 GlyGlnProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGCGCGCTGCGCGAGCGCGCGAGAGCGCGCGAGAGCGCGAGAGCGCGAGAGAG 758
QY 238 GlnMetGlySerArgThrArgAspArgLeuAspGlnValLysGlnGlnGlnGlnGlnGln 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGAGAGCGCGAGCGCGCGAGCGCGCGAGATGAGATGAGATGAGATGAGATGAGATG 818
QY 258 ArgAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CCGACCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
QY 278 ArgLeuLysSerThrPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCTCTAAAGGCTGTTCAGAGCTCTGTGTGAAGACATACGCGCGAGAGAGAGAGAGAGAG 938
QY 298 ValGlnValGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAGAGAGATGAGTGGGCGGTGAGAGATAGCTCTCCACCTCTCTGCGCCAGTGAAT 998

```

RESULT 4
US-08-949-155-51
Sequence 51, Application US/08949155
Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hilder, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ. ID NO.: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-51

```

```

Alignment Scores:
Pred. No.: 2,39e-87 Length: 4267
Score: 994.00 Matches: 217
Percent Similarity: 58.20% Conservative: 42
Best Local Similarity: 48.76% Mismatches: 47
Query Match: 62.56% Indels: 140
DB: 4 Gaps: 5

```

US-09-827-854-16 (1-317) x US-08-949-155-51 (1-4267)

```

QY 1 MetLysValLeuThrPalaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGGCTGTGAAGTGGCTGGT-----CGAGAGAGCGCGAGAGAGAGAG 2489
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArg-----GlnGlnThrGln 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCCGAGG-----CCGCCCGCGAGAGGTGCACGTGTGTGGAGAGAGCCCAAG 2537
QY 38 TTPGInserGlyInArtrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTTP 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGCAGCGAGCCCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2597
QY 58 ValGInThrLeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTCCAGTCCCTGTCTACCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGCACCCAGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGCAACCGTGTCTGTGAGACCTCAGAGCTCCACCGCTCGGGTTCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTTGTGCCAACCTTGTGGGGGTCTGGGTCTGTGTTCTTTTCTTCTCTCTCTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAATTCTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 2896
QY 78 ----- 78

```


Dd	2837	TTGGGAGGAAAAAATCTTTCCTTTCTTTCTTCAATTGACATTCATGTACTTCCTTTCC	28936
Oy	78	-----ArgAlaLeuMetAspGluThrMetIysGluLeuLysAlaTyr	78
Dd	2897	ATCTTGAGCTCTGCCTTCGCTGCTCTCTGGGAGCGACTCTTGCCGCTCCTGCTGTCTGG	29566
Oy	78	-----	78
Dd	2957	AATCTCTGGCACGTCCTCGGCCATTCGCCAGTCACAGACCCTCTTCCTCCCTCACGCC	30166
Oy	79	-----ArgAlaLeuMetAspGluThrMetIysGluLeuLysAlaTyr	92
Dd	3017	CCCGCCCTCTCTTCGCCCCAGAGGAGCTGTATGAGAGAACGATGAAGAGAGTGAAGCCCTAC	30766
Oy	93	LysSerGIuLeuGIuGIuGIuGIuLeuThrProValAlaGIuGIuThrArgAlaIArgIeuSer	112
Dd	3077	GCGAGGAGCTGGAGGCGGACACTGGCCCCCTGATCCACCAAGAACAAGCGAGGCGCTGTCC	31366
Oy	113	LysGIuLeuGIuAlaGIuAlaGIuAlaIArgIeuGIuYalaIAspMetGluAspValCysGIuYar	132
Dd	3137	AAGGAGCTCGAGGCGGCGAGGCCCGCGGTGGCGCCGACATGAGAGAGTGGCCAACGCC	31966
Oy	133	LeuValGIuThrArgGIuGIuValGIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArg	152
Dd	3197	TTCGGTCTTAACCCACGAGGAGTGCACAACTTTGGCGCAACACACGAGAGAGTGGCG	32566
Oy	153	ValArgIeuAlaSerHisIeuArgLysIeuArgLysArgLysArgLysArgAspAlaIAsp	172
Dd	3257	AGCGCCCTGCTTCCACACTCCGCAAGCTGCGCAACAGGCGCTCTCCGACACCGAGAC	33166
Oy	173	LeuGIuLysCysIeuAlaValTyrGIuAlaGIuYalaIArgGIuGIuYalaGIuArgLysLeu	192
Dd	3317	CTGGCAAGAGCCCTGGCGCTGTACTACAGGCGGCGCTGCCGAGGCGCGAGCGCAGCTG	33766
Oy	193	SerAlaIeArGIuArgIeuGIuYProLeuValGIuGIuGIuArgValArgAlaIaIthr	212
Dd	3377	AGCGCCCTCCCGACAGGCTCTCGGCCCCCTGTGTGAGACAGAGGCCAATTGCGCGCCAC	34366
Oy	213	ValGIuSerLeuAlaGIuGIuProLeuGIuGIuArgAlaGIuAlaITRPolYGIuArgLeu	232
Dd	3437	CTGAGTATCACAGGCGCGGCGACGCGTGCAGAGCGCGGAGAGCCTGGGCGCAGAACTG	34966
Oy	233	ArgAlaIArgMetGIuGIuMetGISerATyrArgAspArgIeuAspGluValLysGIu	252
Dd	3497	GCGGAGGAGTGGAGGAGATGGGACGCCGAGCCCGGACCGCTGTGATGATGTGTAG	35566
Oy	253	GlnValAlaGIuValArgAlaLysIeuGIuGIuAlaGIuAlaGIuAlaIleArgIeuGIuAla	272
Dd	3557	CAGCTGGAGGAGGCGCCACCAAAATGGTGTGGACCCCTGTGTGGAAGACATACGGGCG	36166
Oy	273	GluAlaIheGlnAlaIArgLysSerTrpPheGIuProLeuValGIuAspMetGlnArg	292
Dd	3617	GAGGATATTCACGCGCCCTCCCAAAGGCTGTTCGAGCCCTGTGTGGAAGACATACGGGCG	36766
Oy	293	GlnTrpIlaGIuLeuValGIuLysLysGIuAlaIleAlaValGIu---ThrSerAlaIlePro	311
Dd	3677	CAGTGGGCGGCGGCTGTGGAGAGGATGCAGTGCGCCGTGAGCATAAAGCTCTCCACCTCT	37366
Oy	312	ValProSerAspAsn	316
Dd	3737	GCGGCCAGTGATAAT	3751
RESULT 6			
US-08-726-306A-28			
; Sequence 28, Application US/08726306A			
; Patent No. 5958684			
; GENERAL INFORMATION:			
; APPLICANT: van Leeuwen, Frederik Willem			
; APPLICANT: Burbach, Johannes Peter Henri			
; APPLICANT: Grosfeld, Franklin G			
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS			
; NUMBER OF SEQUENCES: 189			
; CORRESPONDENCE ADDRESSES:			

	ADDRESSEE: Banner & Witcoff, Ltd.
	STREET: 1 Financial Center
	CITY: Boston
	STATE: MA
	COUNTRY: US
	ZIP: 02111
	COMPUTER READABLE FORM:
	MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
	COMPUTER: IBM PC compatible
	OPERATING SYSTEM: PC-DOS/MS-DOS
	SOFTWARE: Wordperfect 6.1
	CURRENT APPLICATION DATA:
	APPLICATION NUMBER: US/08/726,306A
	FILING DATE: 02-Oct-1996
	PRIOR APPLICATION DATA:
	APPLICATION NUMBER: GB 95/20080.4
	FILING DATE: 02-Oct-1995
	PRIOR APPLICATION DATA:
	APPLICATION NUMBER: US 60/009,832
	FILING DATE: 01-Jan-1996
	ATTORNEY/AGENT INFORMATION:
	NAME: Williams, Ph.D., Kathleen M.
	REGISTRATION NUMBER: 34,380
	REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
	TELECOMMUNICATION INFORMATION:
	TELEPHONE: (617) 345-9100
	TELEFAX: (617) 345-9111
	INFORMATION FOR SEQ ID NO: 28:
	SEQUENCE CHARACTERISTICS:
	LENGTH: 660 base pairs
	TYPE: nucleic acid
	STRANDEDNESS: double
	TOPOLOGY: linear
	MOLECULE TYPE: cDNA
	US-08-726-306A-28
	Alignment Scores:
	Pred. No.: 1,096-86 Length: 660
	Score: 976.00 Matches: 196
	Percent Similarity: 98.49% Conservative: 0
	Best Local Similarity: 98.49% Mismatches: 3
	Query Match: 61.42% Indels: 0
	DB: 2 Gaps: 0
US-09-827-854-16 (1-317) x US-08-726-306A-28 (1-660)	
OY 1 MettysValleutRPAlaalaLeuDeuValThrPheLeuAlagIcYsgInalalysVal 20	
Dd 62 ATGAAGGTTCTGTGGCGTCCGTCCTGTGCATTTCCGCGCAGATGCCAGCAAGTG 122	
OY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlngInthrGluTPGInsEr 40	
Dd 122 GAGCAAGCGGTGGAGACAGACC CGAGCC CCGACTGCGCCAGCAGACGAGTGGCACAGC 181	
OY 41 GlyGlnArqTTPGiuleuAlaLeuGIYAphrThrpAspTYrlaEuArqTPvaIGInThr 60	
Dd 182 GCCCAAGCCCTGGAACTGGCACTGGCTGGCTTTGGGATTACTGCCCCCTGGGTGCAGACA 241	
OY 61 LeuSerGIuGlnValaGlngIuGluLeuLeuSerGIuValThrGlngIuLeuArGaLa 80	
Dd 242 CTGTCTGAGCAGGTGGACAGAGAGACTGCTCAGTCCCAAGTCACCCAAAGCAATGAGGGCG 301	
OY 81 LeuMetaspGIuThrMetLlysgIuLeuLYsaLaTyrlYssErGIuLeuGIuGlngInLeu 100	
Dd 302 CTGATGGACGAGACCATGTAAGAATTGAAGCCCTAACAAATCGGAACGTGAGGAACAAC TG 361	
OY 101 ThrProVALaGluGluThrArgAlaArqLeuSerLySGluLeuGlnAlaInAla 120	
Dd 362 ACCCGCGTAGGCGAGAGAGACCGCGCAOGCGCTGTCCAAGAGCTGCAGAGCGCGCAGGCC 422	
OY 121 ArgLeuGIyAlaaspMetGIuAspValcYsgIyArqLeuValGlnTyrrArgIyGluVal 140	
Dd 422 CGGCTGGGGCCGGACATGGAGAGAGCTGTGGCGCCCTGGTGGCATGTACCGCGCGAGAGTG 481	

QY 141 GlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 482 CAGCCATGCTCGGCGACACGAGAGCTCGGGTGGCTCGCTCCACCTCCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 542 AAGTGGTGAAGCGCTCCCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGCGAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlyGluArgGlyLeuSerAlaLeuArgGluArgLeu 199
|||||
Db 602 CAGGCGGGGCGCGGAGGCGCGGAGCGGCTCAGCGCCATCCGCGAGCGCTG 658
RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5525493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/849,389
; FILING DATE: 19920519
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/168/DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: HUMAN LIPOPROTEIN E GENE
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..330
; US-07-849-389-6
Alignment Scores:
Pred. No.: 5.47e-41 Length: 330
Score: 503.00 Matches: 106
Percent Similarity: 97.27% Conservative: 1
Best Local Similarity: 96.36% Mismatches: 3
Query Match: 31.66% Indels: 0
Gaps: 0
US-09-827-854-16 (1-317) x US-07-849-389-6 (1-330)
QY 87 LysGlnLeuLysAlaTyrLysSerGluLeuGluGlnLeuThrProValAlaGluGlu 106
|||||
Db 1 AAGGAGTTGAAGGCTTACAAATCGAATCGAGAGCAACTGACCCCGTGGCGAGAGAG 60

QY 107 ThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMet 126
|||||
Db 61 ACCGGGCGACGGCTGTCTCAAGAGAGCTCGAGCGCGGAGGCGCCCGCTGGCGGACATG 120
QY 127 GluAspValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGln 146
|||||
Db 121 GAGGAGGTGGCGCGCGCTGTGTCAGTACCGCGGAGAGGTGCAGGCAATGCTCGGCGAG 180
QY 147 SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeu 166
|||||
Db 181 AGCAGCGAGAGCTGGGGGCGCGCTCCCTCCACCTGCGCAAGCTGCGTAAGCGGCTC 240
QY 167 LeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGlu 186
|||||
Db 241 CTCGCGCATCCCATGATACCTGCGAAGAGCGCTGCGCATGTACCAAGCGCGGCGCGAG 300
QY 187 GlyAlaGluArgGlyLeuSerAlaAlaLeuArg 196
|||||
Db 301 GCGCGGAGCGCGCGCTCAGCGCATCCGC 330
RESULT 8
US-08-617-256-24
; Sequence 24, Application US/08617256
; Patent No. 6043031
; GENERAL INFORMATION:
; APPLICANT: Kyster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,256
; FILING DATE: March 18, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: March 17, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth A.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: SQI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-617-256-24
Alignment Scores:
Pred. No.: 1.43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
Gaps: 0
US-09-827-854-16 (1-317) x US-08-617-256-24 (1-252)

QY 109 AlAArgLeuSerIysGluLeuGlnAlaIaGlnAlaArgLeuGlnIalaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGAGCATGTGAGAGAC 61
QY 129 ValCysGlyArgLeuValAlaGlnIleValGlyGluValGlnAlaMetLeuGlnIleSerThr 148
Db 62 GTGTGGCC-GGCTGTGTGACGTACCGCGGAGGTGCAGGCTCGGCGCAAGACACC 120
QY 149 GlGluLeuArgValArgLeuAlaSerHisLeuArgIysLeuArgIysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCTCGCTCCCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnIysCysLeuAlaValIleGlnAlaGlnIalaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAGTCCCTGTGACGTATACAGGCGCGGCGCGCCGAGAGCGGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCTCTC 252
RESULT 9
US-09-287-141-24
; Sequence 24, Application US/09287141
; Patent No. 6197498
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McCauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,141
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA: 08/406,199
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELEPHONE: (617)227-5941
; TELEFAX: (617)227-7400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-141-24
Alignment Scores:
Pred. No.: 1.43e-28 Length: 252
Score: 374.00 Matches: 82

Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0
US-09-827-854-16 (1-317) x US-09-287-141-24 (1-252)
QY 109 AlAArgLeuSerIysGluLeuGlnAlaIaGlnAlaArgLeuGlnIalaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAGAGCTGCAGAGCGCGCCAGCGCCGCTGGCGCGGAGCATGTGAGAGAC 61
QY 129 ValCysGlyArgLeuValAlaGlnIleValGlyGluValGlnAlaMetLeuGlnIleSerThr 148
Db 62 GTGTGGCC-GGCTGTGTGACGTACCGCGGAGGTGCAGGCTCGGCGCAAGACACC 120
QY 149 GlGluLeuArgValArgLeuAlaSerHisLeuArgIysLeuArgIysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCTCGCTCCCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnIysCysLeuAlaValIleGlnAlaGlnIalaArgGluGlyAla 188
Db 181 GATGCCGATGACCTGCAGAGTCCCTGTGACGTATACAGGCGCGGCGCGCCGAGAGCGGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCTCTC 252
RESULT 10
US-09-431-613-24
; Sequence 24, Application US/09431613
; Patent No. 6221601
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McCauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA: 08/406,199
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELEPHONE: (617)227-5941
; TELEFAX: (617)227-7400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:

Pred. No.:	1,43e-28	Length:	252
Score:	374.00	Matches:	82
Percent Similarity:	97.62%	Conservative:	0
Best Local Similarity:	97.62%	Mismatches:	2
Query Match:	23.54%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-16 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerLySGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGGCGAGCCCGGCTGGCCGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGAGTACGCGCGGCGAGGTGCAGGCGCATGCTCGCGCAGAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTGCCCTCCACCTGCCCAAGCTCGTAAGGCGCTCCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGAGCGGGGCCCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehnman White & McAniff LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504, 245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617, 256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406, 199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:

Pred. No.:	1,43e-28	Length:	252
Score:	374.00	Matches:	82
Percent Similarity:	97.62%	Conservative:	0
Best Local Similarity:	97.62%	Mismatches:	2
Query Match:	23.54%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-16 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerLySGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGGCGAGGTGCAGGCGCATGCTGGCCGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGAGTACGCGCGGCGAGGTGCAGGCGCATGCTGGCCGACAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTGCCCTCCACCTGCCCAAGCTCGTAAGGCGCTCCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGAGCGGGGCCCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehnman White & McAniff LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287, 682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617, 256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406, 199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-682-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGCATGGAGGAGC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyInsThr 148
Db 62 GTGTGGCC-CGCTGTGTGCAATACCCCGCGAGCTCAGGCCATGCTCGGCCAGAGCACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGTGGCGCTCCGCCCTCCACCTGCGCAAGCTGGTAAGGGCTCTCTCCGC 180

QY 169 AspaIaAspAspLeuGlnIysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTGCAGAAATCCCTGCGCATGTACAGCGCGCGCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 1,43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-679-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCAGCCCGCTGGCGCGGCATGGAGGAGC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyInsThr 148
Db 62 GTGTGGCC-CGCTGTGTGCAATACCCCGCGAGCTCAGGCCATGCTCGGCCAGAGCACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGTGGCGCTCCGCCCTCCACCTGCGCAAGCTGGTAAGGGCTCTCTCCGC 180

QY 169 AspaIaAspAspLeuGlnIysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTGCAGAAATCCCTGCGCATGTACAGCGCGCGCGCGAGGGCGCC 240

QY 189 GluArgGlyLeu 192
Db 241 GAGCGGCGCCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

```

: APPLICATION NUMBER: US/09/397,766
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/617,256
: FILING DATE: 18-MAR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 24736-2002I
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-450-8400
: TELEFAX: 619-587-5360
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: US-09-397-766-24

Alignment Scores:
Pred. No.: 1.43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-16 (1-317) x US-09-397-766-24 (1-252)
QY 109 AIAATGLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCACAGAGCTGCAGGCGGCGCGCCGCTGGCGGACATGAGAGAC 61
QY 129 VALCYSGIArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTGCGC-CGCGCTGTGTCAGTACCGCGCGGCGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120
QY 149 GIUGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTGCGCTCGCTCCACCTGCGCAAGCTCGTAAAGCGGCTCCTCCGC 180
QY 169 ASPAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGAGTACCTCAGAGTCTCTGGCAGTGTACCAAGGCGGCGCCGCGAGGGCGCC 240
QY 189 GIuArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 15
US-09-287-681-24
: Sequence 24, Application US/09287681
: Patent No. 6277573
: GENERAL INFORMATION:
: APPLICANT: K ster, Hubert
: TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Heller Ehrman White & McCauliffe
: STREET: 4250 Executive Square, 7th Floor
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037-9103
```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/287,681
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/617,256
: FILING DATE: 18-MAR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 24736-2002F
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-450-8400
: TELEFAX: 619-587-5360
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: US-09-287-681-24

Alignment Scores:
Pred. No.: 1.43e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.54% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-16 (1-317) x US-09-287-681-24 (1-252)
QY 109 AIAATGLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCACAGAGCTGCAGGCGGCGCGCCGCTGGCGGACATGAGAGAC 61
QY 129 VALCYSGIArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GGTGTGCGC-CGCGCTGTGTCAGTACCGCGCGGCGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120
QY 149 GIUGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGCTGCGCTCGCTCCACCTGCGCAAGCTCGTAAAGCGGCTCCTCCGC 180
QY 169 ASPAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGAGTACCTCAGAGTCTCTGGCAGTGTACCAAGGCGGCGCCGCGAGGGCGCC 240
QY 189 GIuArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252
```

Search completed: March 14, 2003, 20:18:40
Job time : 39.7343 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1589
Sequence: 1 MKVLMALLVFLAGCOAKV.....VEKVOAAVGTSAVPSPDNH 317

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09827854/runat_11032003.101612.27578/app.query.fasta.1.3576
-DB=Published_Applications_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=Dlosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=ext -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854 @CCN 1.1.298 @runat_11032003.101612.27578
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAP -LARGEOBERT -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEROUT=120 -WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	1156	10	US-09-827-854-9
2	1582	99.6	1156	10	US-09-827-854-10
3	1577	99.2	1156	9	US-09-870-759-129
4	1577	99.2	1156	9	US-09-802-640-17

5	1577	99.2	1156	10	US-09-827-854-8	Sequence 8, Appl1
6	1577	99.2	1291	12	US-10-044-090-454	Sequence 454, App
7	1573	99.0	1156	10	US-09-827-854-12	Sequence 12, Appl
8	1569	98.7	1156	10	US-09-827-854-11	Sequence 11, Appl
9	1568	98.7	1157	10	US-09-954-456-760	Sequence 760, App
10	1568	98.7	1157	10	US-09-880-107-2244	Sequence 2244, App
11	1565	98.5	1156	10	US-09-827-854-7	Sequence 7, Appl1
12	1368.5	86.1	41907	10	US-09-967-013-5	Sequence 5, Appl1
13	859.5	54.1	786	10	US-09-925-102-133	Sequence 133, App
14	625	39.3	478	10	US-09-964-824-374	Sequence 374, App
15	625	39.3	478	10	US-09-880-107-2491	Sequence 2491, App
16	458.5	28.9	449	10	US-09-960-352-4726	Sequence 4726, App
17	455	28.6	356	10	US-09-960-352-5420	Sequence 5420, App
18	439.5	27.7	442	10	US-09-960-352-9395	Sequence 9395, App
19	435	27.4	414	10	US-09-960-352-4237	Sequence 4237, App
20	428.5	27.0	416	10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.6	423	10	US-09-960-352-14047	Sequence 14047, A
22	413	26.0	253	10	US-09-179-5368-130	Sequence 130, App
23	405.5	25.5	409	10	US-09-960-352-5148	Sequence 5148, App
24	401.5	25.3	416	10	US-09-960-352-9797	Sequence 9797, App
25	387.5	24.4	425	10	US-09-960-352-4497	Sequence 4497, App
26	380.5	23.9	390	10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391	10	US-09-960-352-1196	Sequence 1196, App
28	379.5	23.9	392	10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401	10	US-09-960-352-1187	Sequence 4187, App
30	375.5	23.6	453	10	US-09-960-352-8042	Sequence 8042, App
31	374	23.5	252	10	US-09-796-416-24	Sequence 24, Appl
32	374	23.5	252	10	US-09-879-341-24	Sequence 24, Appl
33	369.5	23.3	404	10	US-09-960-352-2325	Sequence 2325, App
34	369.5	23.3	405	10	US-09-960-352-14063	Sequence 14063, A
35	351.5	22.1	377	10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	353	10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378	10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	386	10	US-09-960-352-10361	Sequence 10361, A
39	348.5	21.9	388	10	US-09-960-352-11966	Sequence 11966, A
40	345.5	21.7	396	10	US-09-960-352-3457	Sequence 3457, App
41	344.5	21.7	377	10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.3	432	10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332	10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.2	376	10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.1	373	10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-827-854-9
Sequence 9, Application US/09827854
Patent No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
TITLE OF INVENTION: Compounds and methods for lowering
TITLE OF INVENTION: cholesterol levels without inducing
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-9
Alignment Scores:
Pred. No.: 3.42e-137 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-16 (1-317) x US-09-827-854-9 (1-1156)

QY 1 MetlysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||
DB 61 ATGAAGCTTCTGGGGCTGGTGTGTGTCACATCTCTGGCAGAGATGCCAGGCCAAGGTC 120

QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
   |||||
DB 121 GACCAAGCGGTGGAGACAGACCGGAGCCGAGACTCGCCAGACAGACCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
DB 181 GGGCAGCGCTGGGAGACTGGGACTGGGTGGTGGGATTACTGGGCTGGGTGGCAGACA 240

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
DB 241 CTGCTGAGCAGGTGACAGAGAGAGTGTCTCAGCTCCAGGTCAACCAGGAACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
   |||||
DB 301 CTGATGACAGACACATGAAGAGGTGAAGGCTTACAAATCGAGACTGGAGAACACTG 360

QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
   |||||
DB 361 ACCCGGTTGGCGGAGAGACCGGGCAGCGCTGTCCAGAGAGTGCAGGGCGGCGAGCC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
   |||||
DB 421 CGGCTGGGGCGGACATGTGGAGCGTGTGGCGCGCTGTGCATACCGCGCGGAGGTG 480

QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgLeuValArgLeuAlaSerHisLeuArg 160
   |||||
DB 481 CAGGCAATGCTGGCCAGACACCGAGAGACTCGGGGTGGCTCCCTCCCTCCAGCTGGCG 540

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
   |||||
DB 541 AAGCTGCGTAAAGCGGCTCTCCCGGATGCCGATGACCTGCAGAAAGTCCGAGGAGTAC 600

QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
   |||||
DB 601 CAGGCGGGGGCCGCGAGGGCGCGGAGCGGCTCAGCGCCATCCGCGAGCGGCTGGGG 660

QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220
   |||||
DB 661 CCCCTGGTGGACAGGGCCCGCTGGCGGGCCGCACTGTGGGCTCCCTGGCGGCGCAGCCG 720

QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
   |||||
DB 721 CTACAGAGAGCGGCGCCAGGCTGGGCGGAGCGGCTCGCGCGCGGATGGAGAGTGGGG 780

QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
   |||||
DB 781 AGCGGAGACCGGACCGCCCTGAGCGAGGTGAAGAGCAGGTGGCGAGGTGGCGCGCAGAG 840

QY 261 LeuGlnGlnGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
   |||||
DB 841 CTGGAGAGAGAGCCCAACGACGATACGCTGCGAGGCGGAGGCTCTTCCAGGCGCCCTCAG 900

QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
   |||||
DB 901 AGCTGTTGGAGCGCTCTGGTGGAGAACATGACAGCGCAGTGGGCGGCTGGTGGAGAG 960

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
   |||||
DB 961 GTGACAGGTGCGGTGGGACACAGCGCGCGCTGTGTCGCCAGCAGCATATCAG 1011

RESULT 2
US-09-827-854-10
; Sequence 10, Application US/09827854
```

```
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-827-854-10

Alignment Scores:
Pred. No.: 1,5e-136 Length: 1156
Score: 1582.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.56% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-16 (1-317) x US-09-827-854-10 (1-1156)

QY 1 MetlysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
   |||||
DB 61 ATGAAGCTTCTGGGGCTGGTGTGTGTCACATCTCTGGCAGAGATGCCAGGCCAAGGTC 120

QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
   |||||
DB 121 GACCAAGCGGTGGAGACAGACCGGAGCCGAGACTCGCCAGACAGACCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
DB 181 GGGCAGCGCTGGGAGACTGGGACTGGGTGGTGGGATTACTGGGCTGGGTGGCAGACA 240

QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
DB 241 CTGCTGAGCAGGTGACAGAGAGAGTGTCTCAGCTCCAGGTCAACCAGGAACTGAGGGCG 300

QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
   |||||
DB 301 CTGATGACAGACACATGAAGAGGTGAAGGCTTACAAATCGAGACTGGAGAACACTG 360

QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
   |||||
DB 361 ACCCGGTTGGCGGAGAGACCGGGCAGCGCTGTCCAGAGAGTGCAGGGCGGCGAGCC 420

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
   |||||
DB 421 CGGCTGGGGCGGACATGTGGAGCGTGTGGCGCGCTGTGCATACCGCGCGGAGGTG 480

QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgLeuValArgLeuAlaSerHisLeuArg 160
   |||||
DB 481 CAGGCAATGCTGGCCAGACACCGAGAGCTGGGGTGGCTCCCTCCCTCCAGCTGGCGC 540

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
   |||||
DB 541 AAGCTGCGTAAAGCGGCTCTCCCGGATGCCGATGACCTGCAGAAAGTCCGAGGAGTAC 600

QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
   |||||
DB 601 CAGGCGGGGGCCGCGAGGGCGGCGAGCGGCTCAGCGCCATCCGCGAGCGGCTGGGG 660

QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220
   |||||
```


Alignment Scores:			
Pred. No.:	4, 31e-136	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	9	Gaps:	0
US-09-827-854-16 (1-317) x US-09-802-640-17 (1-1156)			
QY	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20	
DB	61 ATGAAGGTTCTGTGGGCTGGCTTCTGTCTGCATCTCTGGCAGAGATCCAGGCGCAAGT	120	
QY	21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40	
DB	121 GAGCAAGCGGTGGAGACAGAGACCCGAGCCAGCTGGCCAGCAGACCCGAGTGGCAGAGC	180	
QY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60	
DB	181 GGCCAGCCCTGGGAACCTGGCACTGGGTGGCTTTGGGATTAACCTGGCCCTGGGTGGCAGCA	240	
QY	61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValTrpGlnGluLeuArgAla	80	
DB	241 CTGTGTGACGAGGTGCAGAGAGACTGCTCAGCTCCAGGTCACCCAGAGAACTGAGGGCG	300	
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu	100	
DB	301 CTGATGGCAGACCATGAAAGATTGAAGCCCTCAAAATCGCAACTGGAGAAACAATC	360	
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
DB	361 ACCCGGTGGCGGAGAGACCGGGCACGGCTGTCCAAAGAGCTGCAGGGGCGCAAGCC	420	
QY	121 ArgLeuGlnYAlaAspMetGluAspValCysGlyTrpGluValGlnTrpArgGlyGluVal	140	
DB	421 CGCGTGGGGCGGACATGAGAGACTGTGGCGCCGCTGTGTGCAGTACCGCGGCGAGTG	480	
QY	141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160	
DB	481 CAGGCATGCTCTGGCCAGACACCGAGAGACTGGGGGTGGCCCTCCACCTGGCCG	540	
QY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp	180	
DB	541 AAGCGCCTTAAGCGGCTCTCCGGAGATGCCCATCTCGAGAAAGCCCTCGGACAGTAC	600	
QY	181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly	200	
DB	601 CAGCGCGGGGCGCCGAGAGGCGCGGACGCGGCTCACGCCCATCCGCGAGCCCTGGGG	660	
QY	201 ProLeuValGlnGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlnPro	220	
DB	661 CCCCTGGTGAACAAGGGCGGGCTGGGGCGGCCACTGTGGGCTCCCTGGCGGCGACGG	720	
QY	221 LeuGlnGlnArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly	240	
DB	721 CTACAGGAGCGGGCCAGAGGCTCTGGGGGAGAGGGCTGGCCGCGGATGAGAGATGGCG	780	
QY	241 SerArgTrpArgAspArgLysAspGluValLysGlnGlnValAlaGlnValArgAlaIys	260	
DB	781 AGCCGGACCCCGACCGCTGTGAGAGATGAAGAGGATGGCGGGAGGTGGCGGCCAAG	840	
QY	261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280	
DB	841 CTGAGAGAGAGCGGCCAGAGATACGCTGCAGGCCGAGAGGCTTCCAGGCGCCGCTCAAG	900	
QY	281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300	
DB	901 AGCTGGTTCCAGCCCTGGTGAAGACATCAGACGCCAAGTGGGGCGGTGGTGGAGAAAG	960	
QY	301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317	
DB	961 GTGCAAGCTGCGCTGGGCAACAGCGGCCCTCTGTGCGCAGCAATATCC	1011	

RESULT 5					
US-09-827-854-8					
; Sequence 8, Application US/09827854					
; Patent No. US2002012033A1					
; GENERAL INFORMATION:					
; APPLICANT: Zanis, Vassilis					
; APPLICANT: Kypros, Kyriakos E.					
; TITLE OF INVENTION: Compounds and methods for lowering					
; FILE REFERENCE: 07180/004003					
; CURRENT FILING DATE: 2001-04-05					
; PRIOR APPLICATION NUMBER: US 09/679,088					
; PRIOR FILING DATE: 2000-10-04					
; PRIOR APPLICATION NUMBER: US 09/544,386					
; NUMBER OF SEQ ID NOS: 19					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 8					
; LENGTH: 1156					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-827-854-8					
Alignment Scores:					
Pred. No.:	4,31e-136	Length:	1156		
Score:	1577.00	Matches:	316		
Percent Similarity:	99.68%	Conservative:	0		
Best Local Similarity:	99.68%	Mismatches:	1		
Query Match:	99.24%	Indels:	0		
DB:	10	Gaps:	0		
US-09-827-854-16 (1-317) x US-09-827-854-8 (1-1156)					
OY	1	Methylvalleutrpalaaleuleuvalthrphieleualaglycstinalalysval	20		
Dd	61	ATGAAGGTTCTGTGGGCTGCGTTCGTGCACATTCTCCGCAAGATGCCAGGCAAGGTG	120		
OY	21	GIUGINALAVALGLUTHRGILUPROGIUPLROGIULEUAARGINGLNRHGRTPGINSER	40		
Dd	121	GAGCAGAAGCGTGAGACAAGACCAGGCCGAGCTCCGACAGCACCCAGATGGCAGAGC	180		
OY	41	GLYGINARGTTRPGLULEUALALEUGLYARPHETRPASPTYLEUARGTTPVALGINTHR	60		
Dd	181	GGCCAGCGGTGGAACTGGCACTGGGTGCTTTGGGATTACTCGCGGTGGTGACAGACA	240		
OY	61	LEUSERGLUGINVALIGLNUGLIULEULEUSERSEGINVALITHRTNGIULEUARGLA	80		
Dd	241	CTGTGTGAGCAGGTGCAGAGAGAGCTGCTCACCTCCACAGTCACCCAGGAATGAGCGCG	300		
OY	81	LEUMETASPGLUThMERTLYSGIULEULYSALATYRFLYSSERGLIUENGLUGINLEU	100		
Dd	301	CTGATGACAGACCATGTAAGNATTGAAGGCTTACAAATTCGAACTGAGAGAACAACTG	360		
OY	101	THIRPROVALAAGLUGLUThRARGLAARGLEUSERLYSGIULEUNGINALAAGINALA	120		
Dd	361	ACCCGGTGGCGGAGAGACCGCGGACCGGCTGTCCAAGAGACTGCAGCGGCGCAGGCC	420		
OY	121	ARGLEUGLYALAASPMECLIASPYALCYSGIYAARGLEUVALGINTYRRAGLGIVUAL	140		
Dd	421	CGGCTGGGCGCGCATGTGAGAGACGTGTCCGCCCTGCTGTGACGTACCGCGGAGAGTG	480		
OY	141	GINALAMETLEUGLYGLSERTHRGUNGLULEUARVALARGLEUALASERHTISLEUAR	160		
Dd	481	CAGGCCATGCTTGGCCACAGACACCGAGAGCTGCGGGTCCGCTCCCTCCACTCGCCG	540		
OY	161	LYSLEUARGLYARGLUEUARQSPALAAPSPBLEUGLINLYSCYSTLEUALAVALTYR	180		
Dd	541	AAGCTGCGTAGCGGCTCTCCGCCATGCCATGACTCGACAAGCGGCTGGCAGGTGAC	600		
OY	181	GINALAGLYALARGLUNGLYALAGLUARGLYLEUSERALILEARGLUARGLEUGC	200		

Db 601 CAGGCCGGGGCCCGCAGGGCCGCTCAGCCCATCCGCGAGCCCTGGG 660
Qy 201 ProLeuValIgluInglYargValArAlaIaIthrValIglYSerLeuAlaIgluInpro 220
Db 661 CCCCTGCTGGAACAGGGCCGGCTGGGGCCGACTGTGGCTCCCTGGCCGCGCAGCG 720
Qy 221 LeuIngluArAlaIgluInlaIthrIglYuarIgluArAlaIargMetIgluIumetIly 240
Db 721 CTACAGAGACGGGGCCGAGCCCTGGGGCGAGCGCTGCCGCCGAGATGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspIgluValIlySgluIngluInValAlaIgluValArAlaIys 260
Db 781 ACCCGAGCCCGGACCGCTGGAGCAGGTGAAGAGAGAGTGGCGAGGTGGCGCCGCAAG 840
Qy 261 LeuIgluIngluInlaIgluInIleArIgluInlaIgluInlaIphoIgluInlaIargLeuIys 280
Db 841 CTGGAGAGACAGCCCGCAGCATAGCTCGAGGCCGAGGCTTCCAGGCCCGCCCTCAG 900
Qy 281 SerTrpPheIgluProLeuValIgluIuAspMetIgluArIgluInTrpAlaIgluValIgluIys 300
Db 901 ACCTGGTTCGAGCCCTGGTGGAGACATGCAGCGCATGGGGCGGGCTGGTGAGAG 960
Qy 301 ValIgluInlaIalaIvalIglYThrSerAlaIaIproValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCCGTGGGCAACGCGCCCTGTGCCAGCAGCATCAC 1011
RESULT 6
US-10-044-090-454
; Sequence 454, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 454
; LENGTH: 1291
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 251498CB1
; LOCATION: 46
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-454
Alignment Scores:
Pred. No.: 4,92e-136 Length: 1291
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 12 Gaps: 0
US-09-827-854-16 (1-317) x US-10-044-090-454 (1-1291)
Qy 1 MetIysValLeuTrpAlaIalaIleuValIthrPheLeuAlaIglYcysIgluInlaIalysVal 20
Db 186 ATGAAGGTTCTGTGGCTGCTGGTGTGCATCTCTGGCGAGATGCCAGGCCAAAGGTG 245
Qy 21 GluInlaIalaIvalIgluInThIgluProIgluIuLeuArIgluIngluInThIgluInTrpIgluInser 40
Db 246 GAGCAGAGCGGTGAGACAGACGCGGAGCCGAGCTGCGCCACACACACAGCAGTGGCAGAGC 305
Qy 41 GlyIlnArIgluInlaIleuAlaIleuIglYArgPheTrpAspTyrLeuArIgluInTrpValIgluInThr 60
Db 306 GGCCAGCGCTGGAACTGCGACTGGGTCCCTTTTGGGATTTACTCTGCGCTGGTGACAGACA 365
Qy 61 LeuSerIgluInlaIgluIngluInleuLeuSerSerIgluInlaIthrIgluInleuArAla 80

Db 366 CTGCTGACAGCTGACAGAGAGAGCTGCTACACTCCACAGTACCAGCACTGAGAGGGG 425
Qy 81 LeuMetAspIgluThrMetIlySgluLeuIlyAsIaIlyTrIysSerIgluIngluInleu 100
Db 426 CTGATGAGACACCACTGAAGAGTTGAAGGCTTCAAAATCGAACTGGAGACAACTG 485
Qy 101 ThrProValAlaIgluInThrArAlaIargLeuSerIlySgluIleuInlaIalaIgluInla 120
Db 486 ACCCGGTGGGAGAGAGAGCGGGGACAGCGCTGTCCAAAGAGACTCCAGCGCGGCGAGGC 545
Qy 121 ArIgluInlaIalaIspMetIluAspValIlySglYArIgluInlaIgluInTrpArIgluInlaI 140
Db 546 CGGCTGGGGCGGACATGAGAGACGTGTGCCGCCCTGTGTCAGTACCGCGGGCGAGGTG 605
Qy 141 GluAlaMetLeuIgluInSerThrIgluInleuArIgluInlaIaIaIserHisIleuArI 160
Db 606 CAGGCCATGCTGGCCAGACACGAGAGCTGGGGGTGGCTCGCTCCACACTGGCC 665
Qy 161 LysLeuArIglYArgLeuLeuArIgluAspAlaAspAspLeuIgluInIlycysLeuAlaIalYtyr 180
Db 666 AAGCTGCGTAACGGGCTCTCCGCGATGCCGATGACCTGCAGAACGCGCTGGCAGTGTAC 725
Qy 181 GluAlaIglYAlaIargIgluInlaIgluInlaIgluInlaIleuArIgluInlaIgluInly 200
Db 726 CAGGCCGGGGCCCGCAGAGGCGCCAGCGCGCTCAGCGCCATCCGAGCGAGCGCTGGGG 785
Qy 201 ProLeuValIgluInglYargValArAlaIaIthrValIglYSerLeuAlaIgluInpro 220
Db 786 CCCCTGGTGGAAACAGGGCCGTGGGGCCGCTGAGTGGTCCCTGGCCGGCCAGCGG 845
Qy 221 LeuIngluArAlaIgluInlaIthrIglYIgluArIgluArAlaIargMetIgluIumetIly 240
Db 846 CTACAGAGACGGGGCCAGGCTCGGGGCGAGCGGCTGCCGCGCATGGAGAGATGGGC 905
Qy 241 SerArgThrArgAspArgLeuAspIgluValIlySgluInlaIalaIgluValArAlaIys 260
Db 906 ACCCGAGCCCGGACCGCTGGAGCAGGTGAAGAGAGAGTGGCGAGGTGGCGCCCAAG 965
Qy 261 LeuIgluIngluInlaIgluInIleArIgluInlaIgluInlaIphoIgluInlaIargLeuIys 280
Db 966 CTGGAGAGACAGCCCGCAGCATAGCTGCGAGGCCGAGGCTTCCAGGCCCGCCCTCAG 1025
Qy 281 SerTrpPheIgluProLeuValIgluIuAspMetIgluArIgluInTrpAlaIgluValIgluIys 300
Db 1026 ACCTGGTTCGAGCCCTGGTGGAGACATGCAGCGCATGGGGCGGGCTGGTGAGAG 1085
Qy 301 ValIgluInlaIalaIvalIglYThrSerAlaIaIproValProSerAspAsnHis 317
Db 1086 GTGCAGGCTGCCGTGGGCAACGCGCCCTGTGCCAGCAGCATCAC 1136
RESULT 7
US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilius
; TITLE OF INVENTION: Kypreos, Kyriakis E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; FILE REFERENCE: 2001-04-05
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens

Db 336 ATGAAGTTCTGTGGCGCTGGTCTGTCACATTCTGTGGAGAGATGCCAGGCAAGGTG 337
Qy 21 GtuglnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGACGCCGAGCTGCCTCCAGCAGACCGAATGGCAGAGC 277
Qy 41 GtuglnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCCCTGGAACTGGCACTGGTCTTTGGGATTTACCTGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValGlnGluGluLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCGAGCTCCAGGTCAACCCAGAACTGAGGCG 157
Qy 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
|||||
Db 156 GCTGATGGACAGACCAATGAAGAGTTGAAGCCCTACAAATCGGAACGTGAGAACT 97
Qy 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
|||||
Db 96 GACCCCGGTGGCGGAGAGACGGCGGCTGTCCAAAGAGCTGCAGCGCGCGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

RESULT 15

US-09-880-107-2491/C
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-NO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIORITY FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIORITY FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIORITY FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:

Pred. No.: 2.72e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 39.33% Indels: 1
DB: 10 Gaps: 0

US-09-827-854-16 (1-317) x US-09-880-107-2491 (1-478)

Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 336 ATGAAGTTCTGTGGCGCTGGTCTGTCACATTCTGTGGAGAGATGCCAGGCAAGGTG 337
Qy 21 GtuglnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGACGCCGAGCTGCCTCCAGCAGACCGAATGGCAGAGC 277

Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCCCTGGAACTGGCACTGGTCTTTGGGATTTACCTGCTGGGTGCAGACA 217
Qy 61 LeuSerGluGlnValGlnGluGluLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCGAGCTCCAGGTCAACCCAGAACTGAGGCG 157
Qy 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLe 100
|||||
Db 156 GCTGATGGACAGACCAATGAAGAGTTGAAGCCCTACAAATCGGAACGTGAGAACT 97
Qy 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
|||||
Db 96 GACCCCGGTGGCGGAGAGACGGCGGCTGTCCAAAGAGCTGCAGCGCGCGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGGACATGAGAGAGCTG 9

Search completed: March 15, 2003, 03:04:20
Job time : 79.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-16

Perfect score: 1889

Sequence: 1 MKVLMALVTFLLAGCOAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_plus_p2n.model -DEV=xlp  
-O=/cgn2.1/USPFO_spool/US09827854/runat_11032003_101610_27495/app_query.fasta.1.3576  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09827854_gcgn.1.1.8826_ernunt.11032003_101610_27495 -NCPU=6 -ICPU=3  
-NO_XLPRX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*  
1: em_estba:*  
2: em_esthum:*  
3: em_estlin:*  
4: em_estmu:*  
5: em_estov:*  
6: em_estpl:*  
7: em_estro:*  
8: em_hic:*  
9: gb_est1:*  
10: gb_est2:*  
11: gb_hic:*  
12: gb_est3:*  
13: gb_est4:*  
14: gb_est5:*  
15: em_estfun:*  
16: em_estom:*  
17: gb_gss:*  
18: em_gss_hum:*  
19: em_gss_inv:*  
20: em_gss_pln:*  
21: em_gss_vit:*  
22: em_gss_fun:*  
23: em_gss_mam:*  
24: em_gss_mus:*  
25: em_gss_other:*  
26: em_gss_pro:*  
27: em_gss_rtd:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282.5	80.7	1027	13	B1670367
2	1246	78.4	842	12	B6763371
3	1240	78.0	933	13	B1668318
4	1212	76.3	938	12	B6761746
5	1209.5	76.1	942	13	B1600906
6	1195	75.2	800	13	BM042094
7	1194.5	75.2	927	12	BG472299
8	1191.5	75.0	922	13	B1597743
9	1182.5	74.4	817	12	BG774871
10	1181	74.3	811	13	B1600563
11	1167.5	73.5	845	12	BG829472
12	1144	72.0	790	12	BG707147
13	1141	71.8	919	13	B1551475
14	1131	71.2	757	13	BM042228
15	1121.5	70.6	907	12	BG706129
16	1120	70.5	706	14	BM728696
17	1120	70.5	741	12	BG762924
18	1118	70.4	1110	11	AK010261
19	1116	70.2	804	12	BG702752
20	1115	70.2	834	13	BM042676
21	1112.5	70.0	808	13	B1668329
22	1109	69.8	855	13	B1616362
23	1106	69.6	803	13	B1670350
24	1103	69.4	782	12	BG716776
25	1103	69.4	812	13	B1601551
26	1103	69.4	965	14	B0677266
27	1100	69.2	757	13	B1603658
28	1097.5	69.1	797	12	BG715366
29	1097	69.0	846	13	B1159757
30	1096	69.0	757	13	BM042153
31	1091	68.7	757	12	BG707750
32	1091	68.7	794	12	B1601279
33	1091	68.7	954	13	BE967543
34	1089	68.5	790	13	B1551066
35	1088	68.5	790	13	B1551811
36	1084.5	68.3	798	12	BG708414
37	1083.5	68.2	891	13	B1549292
38	1077.5	67.8	802	13	B1458355
39	1076	67.7	1100	14	BM914382
40	1075.5	67.7	914	13	B1603523
41	1075	67.7	748	13	B1553085
42	1074.5	67.6	930	12	BF967857
43	1070.5	67.4	812	12	BG769968
44	1070	67.3	888	13	B1544886
45	1069.5	67.3	803	12	BG709360

ALIGNMENTS

```
RESULT 1  
B1670367  
LOCUS 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',  
DEFINITION B1670367 mRNA sequence.  
ACCESSION B1670367  
VERSION B1670367.1 GI:15584600  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1027)  
NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE
```


/note="Organ: skin; Vector: pOTB7; Site-1: XhoI; Site-2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library." "

BASE COUNT 154 a 250 c 337 g 101 t

Alignment Scores:

Pred. No.: 3.37e-121
Score: 1246.00
Percent Similarity: 96.67%
Best Local Similarity: 95.93%
Query Match: 78.41%
DB: 12

Length: 842
Matches: 259
Conservative: 2
Mismatch: 8
Indels: 2
Gaps: 0

US-09-827-854-16 (1-317) x BG763371 (1-842)

Qy 1 MetlysalVallEutrpAlaAlaLeuValThrpheLeuAlaGlyCysGlnAlaVal 20
Db ATGAAGGTTCTGTGGCGCTTGTGTGTCATTCTGCGAGGATGCCAAGGTG 95
Qy 21 GluGlnAlaValAGluThrpGluProGluProGluLeuArgGlnGlnThrpGlnSer 40
Db GAGCAAGGGGTGGAGACAGACCGGAGCCGAGCTGCCAGACAGCCGATGGCAGAGC 155
Qy 41 GlyGlnArGTrpGluLeuAlaLeuGlyArpPheTrpAspTyrLeuArgTrpValGlnThr 60
Db GGCCAGCCCTGGGAACCTGGCACTGGGTCCTTTGGGATTAACGCGCTGGGTGAGACA 215
Qy 61 LeuSerGluGlnValAGluGlnGluLeuLeuSerSerGlnValThrpGlnLeuArgAla 80
Db CTGCTGTGAGCAGGTGACAGAGAGCTGCTCCAGCTCCAGCCAGAACCTGAGGCG 275
Qy 81 LeuMetAspGluThrpMetlysalVallEutrpAlaAlaLeuValThrpheLeuAlaGlyCysGlnAlaVal 100
Db CTGATGACGAGACCATTAAGAGAGTGAAGGCTTCAATGAGGAGAGCACTGAGCACTG 335
Qy 101 ThrProValAlaGluGluThrpArGAlaArGLeuSerlyGluLeuGlnAlaAlaGlnAla 120
Db ACCCGGGGGGAGAGAGACCGGGGCAAGGCTGTCCAGAGAGCTGCAAGCGCGAGCGC 395
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArGLeuValGlnTyrArgGlyGluVal 140
Db CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCTGTCAGTACCCGCGAGAGTG 455
Qy 141 GlnAlaMetLeuGlyGlnSerThrpGluLeuArgValArGLeuAlaSerHisLeuArg 160
Db CAGGCAATGCTGGCCAGAGACCGAGAGAGCTGGGCTGCTGCTCCACCTGCGC 515
Qy 161 LysLeuArGlyArGLeuLeuArGAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
Db AAGCTGGCTAAGCGGCTCTCCGCGATGACCTGCGAGAGCGGCTGCGAGTAC 575
Qy 181 GlnAlaGlyAlaArGLeuGlyAlaGlnArGLeuSerAlaIleArgGluArgLeuGly 200
Db CAGGCGGGGGCGGAGAGCGCGGCGGCTGACGCGCTGAGCCGAGCGCTGGGG 635
Qy 201 ProLeuValAGluGlnGlyArGValArGAlaAlaThrpValAGlySerLeuAlaGlyGlnPro 220
Db CCCTG-GTGGAGACAGGGCGGCTGGGCGCAACTGTGGCTCCCTGGCGGAGCGG 694
Qy 221 LeuGlnAlaValAGlnAlaTrpGlyGluArGLeuArgAlaArGMetGluGlnMet-G1 240
Db CTACAGAGAGGGGGCGGAGCGCTGGGGGAGGCGCTGCGCGGCGGATGAGAGATGGG 754
Qy 240 ySerArGTrpArGAspArGLeuAspGluValAGlyGlnGlnAlaGlnValArGAlaVal 260
Db CAGCCGAGCCCGCAGCTGTGAGAGAGTGAAGGAGACAGTGGCGGAGTGGCGGCAAG 814

Qy 260 sleuGluGlnAlaGlnGlnIleArg 269
Db 815 CTGGAGAGCAGCAGCGCCAGATACGC 842

RESULT 3
B1668318
LOCUS
DEFINITION
B1668318 933 bp mRNA linear EST 12-SEP-2001
603295681.F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5',
mRNA sequence.

ACCESSION
B1668318
VERSION
B1668318.1 GI:15582551
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 933)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LHAM1798 row: a column: 21
High quality sequence stop: 780.

FEATURES

source

Location/Qualifiers
1..933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5314844"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site-1: BamHI; Site-2: SalI-XhoI (GTCGAG
); Oligo-dT primed using priming 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 284 c 359 g 119 t

Alignment Scores:

Pred. No.: 1.68e-120
Score: 1240.00
Percent Similarity: 93.73%
Best Local Similarity: 91.29%
Query Match: 78.04%
DB: 13

Length: 933
Matches: 262
Conservative: 7
Mismatch: 15
Indels: 4
Gaps: 1

US-09-827-854-16 (1-317) x B1668318 (1-933)

Qy 1 MetlysalVallEutrpAlaAlaLeuValThrpheLeuAlaGlyCysGlnAlaVal 20
Db ATGAAGGTTCTGTGGCGCTTGTGTGTCATTCTGCGAGGATGCCAAGGTG 134
Qy 21 GluGlnAlaValAGluThrpGluProGluProGluLeuArgGlnGlnThrpGlnSer 40
Db GAGCAAGGGGTGGAGACAGACCGGAGCCGAGCTGCCAGACAGCCGATGGCAGAGC 194
Qy 41 GlyGlnArGTrpGluLeuAlaLeuGlyArpPheTrpAspTyrLeuArgTrpValGlnThr 60
Db CAGCCGAGCCCGCAGCTGTGAGAGAGTGAAGGAGACAGTGGCGGAGTGGCGGCAAG 814

Db	195	GGCCAGCGCTGGGAAGTGGCACTGGGTGGGCTCTTTGGGATTACCTGGGCTGGGTCAGAGACA	254
QY	61	LeuSerGIuGlnValGIuGlnGIuLeuLeuSerSerGIuValThGIuGIuLeuArgAla	80
Db	255	CTGTGTGAGACAGTGCAGAGAGAGAGCTGTCTACACTCTCCACAGTCCAGGACACCGAGACGTAGAGGGCG	314
QY	81	LeuMetAspGIuThMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuGIuLeu	100
Db	315	CTGATGTGACGACACCATGTGAAGAGTGTGAAGGCGCTTCAAAATCGGAATGTGAAGGAACAACATG	374
QY	101	ThrProValAlaGIuGIuThArgAlaArgLeuSerLysGIuLeuGIuAlaAlaGIuAla	120
Db	375	ACCCGGGTGGCGGAGAGACGCGGGGACGCGTGTCCAAAGACACTGCAGCGGGCGACAGGCC	434
QY	121	ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGIuTyrArgGIuVal	140
Db	435	CGGCTGGGGCGGACATGTGAGAGACGTGTGGCGCCGCTGTGCAGTACCGCGCGAGGTG	494
QY	141	GIuAlaMetLeuGIuGIuSerThrGIuLeuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	495	CAGGCCATGCTCTGGCCAGAGACACCGAGAGCTGGGGGTGGCTGCTGCCCTCCACCTGGCGC	554
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysCysLeuAlaValTyr	180
Db	555	AAGCTGCGTAAAGCGGCTCTCCGCGATGGCCATGTACCTGCACAGAAGCGCTGGCAGATGAC	614
QY	181	GIuAlaGIuAlaArgGIuGIuValaGIuArgGIuLysSerAlaIleArgGIuArgLeuGIu	200
Db	615	CAGGCGGGGGCGCGGAGGGCGCCGAGCGCGGCTCTACGCCATCCGCGACGCGCTGGGG	674
QY	201	ProLeuValGIuGIuGIuValArgValArgAlaIleThrValGIuSerLeuAlaGIuIlePro	220
Db	675	CCCTGGTGTGAACAGGGCGCGTGGCGGGCGCCACTGTGGTCTCCCTGGCGGCGACGGC	734
QY	221	LeuGIuGIuArgAlaGIuAlaIleArgLysGIuArgLeuArgAlaArg-MetGIu---GIuMet	239
Db	735	CTACAGAGACGGGGCGCCAGGCTGTGGGGCGAGCGGCTGCGCGCCGATTTGGAGGCACATTG	794
QY	239	tGIuSerArgThrArgAspArgLeuAsp-GluValLysGIuGIuValaAlaGIuValaArga	259
Db	795	GGGCAAGCGCGAGCCCGGAACGCTGTGAAGCAGAGTGAAGGACGAGGCGGAGATTGGCCG	854
QY	259	IaLysLeuGIuGIuGIuGIuAlaGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAlaArgL	279
Db	855	GCCAAAGGCTGGGGGACCAAGGCCCGCAATTCGTTGCAGGCGCAA-GCCCTTCAGAGCGGCT	913
QY	279	eulYsserTrpPheGIu 284	
Db	914	CCAAAGAGCTGCTTCAAA 930	
RESULT 4			
LOCUS	BG761746	938 bp	mRNA linear EST 15-MAY-2001
DEFINITION	6027717942P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:484141 5',		
		mRNA sequence.	
ACCESSION	BG761746		
VERSION	BG761746.1	GI:14072399	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 938)		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: ATCC/DCIT/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be		

FEATURES	source
found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LUCM1674 row: c column: 12 High quality sequence stop: 767. Location/Qualifiers	1. .938 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_id="NIH_MGC_49" /tissue_type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOMB7; site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCGAG(g). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT	171 a 272 c 374 g 121 t
ORIGIN	
Alignment Scores:	
Score: No.:	1.52e-117 938
Percent Similarity:	1212.00 Matches: 271
Best Local Similarity:	92.28% Conservative: 4
Query Match:	90.94% Mismatches: 12
DB:	76.27% Indels: 11
	Gaps: 2
US-09-827-854-16 (1-317) x BG761746 (1-938)	
QY 1 MettValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal	20
Db 57 ATGAAGTCTCTGTGGCGCTGCTGCTGCATTCCTCGGAGAGATGCCAGCCAGGTG	116
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db 117 GAGCAAGCGGTGGAGACAGACCAGGCCCAAGCTGCGCCAGCAGACCAGTGGCAGACG	176
QY 41 GlyGlnArgTrpGluLeuAlaLeuGluLysArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db 177 GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGGATTCCTGGCTGGGTGCAGCA	236
QY 61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db 237 CTGCTGTGACGAGGTGCAGGAGAGACTGCTCAAGTCCACCGACACCCAGAACTGAGGGC	296
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu	100
Db 297 CTGATGACGAGACCATGAAGAGTTGAAGGCTTCAACAATCGAACTGGAGAACCACTG	356
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db 357 ACCCGGGGGGAGGAGACCGGGCAGGCTGTCCAAAGAGCTGCAGGCGGCGCAGGCC	416
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal	140
Db 417 CGGCTGGGGCGGACATGGAGAGACTGTGGCGGCCCTGTGTGCACTACCGCGGAGAGTG	476
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db 477 CAGGCAATGCTCGGCGACAGCACCGAGAGAGTGGGGGTGGCCCTGCCACTGGCC	536
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180
Db 537 AAGTCGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAGACCGCTGGCAGTGTAC	596
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db 597 CAGGCGGGGGCGCGAGGCGCGAGGCGCTGCTCAAGCGCATTCGCGACCGCTGGGG	656

QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySer-LeuAlaGlyGlnPro 220
|||||
Db 657 CCCCTGTGTGAACAGAGGGCGGGCGGGCGGCACACTGTGGCTCCCTTGGCCGGCCAGGC 716
QY 220 OleuGlnGluArgAlaGlnAlaTrp-GlyGluArg-LeuArgAlaArgMetGlu---Glu 238
|||||
Db 717 GCTACAGAGAGCGGGCCAGCGCTGGGGCGGACCGCTTGGCGCGGATGAGAGGACAT 776
QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValLys-GluGlnValAla---Glu 256
|||||
Db 777 TGGGACAGCCGGACCCGGCGACCGCTTGGACGAGGTGAAGGAGCAGGTGGCGGAGC 836
QY 257 ValArgAlaLysLeuGlnGlu-GlnAlaGlnGlnIleArgLeuGlnAla--GlnAlaPhe 275
|||||
Db 837 TTGGGGCCCAACACTGTGAGAGACAGCGCCAGAGATACCCCTCGACGCCGGAAGGCTTC 896
QY 276 GlnAlaArgLeuLysSerTrpPheGluProLeuVal 287
|||||
Db 897 CCAGGCGCGCTCAAAAGCTGTTCGAAAGCCCTGT 932
RESULT 5
B1600906 B1600906 942 bp mRNA linear EST 07-SEP-2001
LOCUS 603249241F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
DEFINITION mRNA sequence.
ACCESSION B1600906
VERSION B1600906.1 GI:15493845
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 942)
NIH-MGC http://mgi.mc.man.ac.uk/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cga@stratford.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshitsugu and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L16M1762 row: a column: 11
High quality sequence stop: 762.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5301010"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to R0.5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 177 a 284 c 361 g 120 t
ORIGIN
Alignment Scores: 2.81e-117 Length: 942
Pred. No.: 1209.50 Matches: 265
Score: 93.13% Conservative: 6
Percent Similarity: 93.13%

Best Local Similarity: 91.07% Mismatches: 14
Query Match: 76.12% Indels: 8
DB: 13 Gaps: 1
US-09-827-854-16 (1-317) x B1600906 (1-942)
QY 1 MetLysValLeuTrpAlaAlaLeuValTrpPheLeuAlaGlyCysGlnAlaLysVal 20
Db 73 ATGAAGGTTCTGTGGGCTCGTTGGTGTGTCACATTCCTGGCAGGATGCCAGGCCAAGGTG 132
QY 21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGCCGGAGCCGAGCTCCGCGACAGACCCAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
Db 193 GGCACGCGCTGGGAACGTGCACTGGGTCTTGGGATTTACTGTGGGTGGGAGACA 252
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValTrpGlnGlnLeuArgAla 80
|||||
Db 253 CTGTCTGACAGAGTGCAGAGAGAGCTCTCAGCTCCAGTCCACCAAGAACTGAGAGCGG 312
QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
|||||
Db 313 CTGATGAGAGAGACCATGAAAGAGTTGAAGGCTTCAAAATCGAACTGGAGGAACAACATG 372
QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGAGAGAGAGCGGGCAGCGCTGTCCAAAGACTCGACGCGCGCAGCGC 432
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
|||||
Db 433 CGCGTGGCGCGGACATGAGAGAGCTGTGCGCCCTGTGTGCTAGTACCGCGGACAGGTG 492
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCGATCTCGGCGCAGAGCACCAGAGAGTGGGGGTGCGCTGCCCTCCACCTGGCGC 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
|||||
Db 553 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGATGACCAAGCGCTGACAGGTGAC 612
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGGCGCGGCGCGGAGAGGCGCCAGCGCGGCTCAGGCGCATCCGACGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGTGAACAGAGGGCGCGTGGCGGC-GCCACTGTGGGCTCCTTGGCGCGCAGCGC 731
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAla-ArgMetGlu-GluMetG 240
|||||
Db 732 CTACAGAGAGCGGCGGAGCGCTGGGGCGAGCGGGTCCGCGCGCGGATGAGGAGATTG 791
QY 240 LysArgTrpArgAspArgLeuAsp-GluValLys-GluGlnAlaAlaGluVal-ArgA 259
|||||
Db 792 GCAAGCGGAGCCGCGACGCGCTGGACCGAGGTGAACGACGAGGTGGCGAAGTGGCGG 851
QY 259 LysLeuGlnGluGlnAla---GlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 278
|||||
Db 852 CCAACTTGGAGAAACAGCGCCAGCAGATATACCTTGCAGGCGGAGGCTTCAGAGGCC 911
QY 278 rLeuLeuSerTrpPheGluPro 285
|||||
Db 912 GCTCTCAAAAC-TGGTTGACCCCT 933
RESULT 6
BM042094 800 bp mRNA linear EST 07-NOV-2001
LOCUS BM042094
DEFINITION 603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
ACCESSION BM042094
VERSION BM042094.1 GI:16771361

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
Tissue Procurement: DCTP/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM1875 row: 1 column: 03
High quality sequence stop: 792.

FEATURES
source
1..800
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_lib="NIH-MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT87; Site_1: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC library."

BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN

Alignment Scores:

Pred. No.: 7.54e-116 Length: 800
Score: 1195.00 Matches: 243
Percent Similarity: 98.808 Conserved: 4
Best Local Similarity: 97.208 Mismatches: 3
Query Match: 75.208 Indels: 1
DB: 13 Gaps: 0

US-09-827-854-16 (1-317) x BM042094 (1-800)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 50 ATGAAGATTCTGTGGCTGCTGTGTCACATTCCTCGCAGAGATGCCAGGCCAAGGTG 109
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnProGlnSer 40
|||||
DB 110 GACCAAGCGGTGAGACAGAGCCGAGCCGACCTCGCCAGACCGAGCGGTGGCAGAGC 169
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 170 GGCACGCGGTGAGACTGCGACTGGGTCTTTGGGATTACTGCGCTGGGTCAGACA 229
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 230 CTGTCTGACAGAGTGCAGAGGAGCTGCTCAGCTCCAGGTCAACCAGAACTGAGGGCG 289
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
|||||
DB 290 CTGATGACAGAGCCATGGAAGGAGTGAAGGCTTACAAATCGAAGTGGAGAACTG 349
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
|||||
DB 350 ACCCCGCTGGCGAGAGACGCGGCGCTGTCCAGAGAGCTGCAAGCGCGCGCAGGCC 409

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 410 CGCGTGGCGCCGACATGAGGAGCGTGTGGCGCGCTGTGACATACCGCGCGAGGTG 469
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 470 CAGGCCATGCTCGGCCAGAGACCGAGAGACTCGGGTGGCTCCCTCCCACTCGCCG 529
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 530 AACCTGCTAAGCGGCTCTCCCGATGCCGATGACCTGCAGAAAGCGCTCGCAGGTGAC 589
QY 181 GlnAlaGlyAlaArgGlnGlnGlyAlaArgGlyLeuSerAlaLeuArgGluGlnGly 200
|||||
DB 590 CAGCGCGGGCGCCGACAGGCGCGGAGCGCGCTCAGCCCTCCGCGAGCGCTGGGG 649
QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
|||||
DB 650 CCCCTGTTGACAGAGGCGCGTGGCGCCAGCACTGTGGCTCCCTGGCGGCCAGCGG 709
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
|||||
DB 710 CTACAGAGAGCGGCGCCAGGCTGGGCGAGCGGTG-CGCGCGGATGAGAGATGGCG 768
QY 241 SerArgThrArgAspArgLeuAspGluVal 250
DB 769 ACCGGGAGCCGCGACGCTGAGCAGCGTG 798

RESULT 7 927 bp mRNA linear EST 21-MAR-2001
BG472299
LOCUS 602513830F1 NIH-MGC_16 Homo sapiens CDNA clone IMAGE:4645759 5',
DEFINITION mRNA sequence.
ACCESSION BG472299
VERSION BG472299.1 GI:13404485
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 927)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LNCM1419 row: k column: 08
High quality sequence stop: 848.

FEATURES
source

1..927
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_lib="NIH-MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOT87; Site_1: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."

BASE COUNT 170 a 271 c 371 g 115 t
ORIGIN


```
Db 313 CTGATGACGACGACATGAAGAGTTGAGGCTTACAAATCGAAGCTGGAGGACAACTG 372
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 373 ACCCGGATGGCGAGGAGGAGCGGCGACGGCTGTCCAAAGAGCTGCAGGGCGGCGAGCC 432
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 433 CGGCTGGCGCGGACATGAGGAGGTGTGGCGCGCTGTGTGACATGCCGGCGAGAGTG 492
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 493 CAGGCCATGCTCGGCCACAGACACGAGAGCTCGGGTGGCGCTCCGCTCCACCTGCGC 552
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 553 AAGCTGGTAAAGCGGCTCTCCGCGATGGCGATGACCTGCAGAAAGCGCGCTGGCAGTGTAC 612
Qy 181 GlnAlaGlyAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluAlaArgLeuGly 200
Db 613 CAGGCGGGGCGCGCAGAGGCGCGGAGCGCGGCTC-AGCCCATCCGCGAGCGCGCTGGGG 671
Qy 201 ProLeuVal-GluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln-P 220
Db 672 CCCCTGGTGGGAACAGAGCGCGGGTGGCGCGCGCAGCTGTGGCTCCCTGGCGGCGACGC 731
Qy 220 roLeuGlnGluArg-AlaGlnAlaTatPglGlyGluArgLeuArgAlaArgMetGluGlnMet 239
Db 732 CGCTAGAGAGAGCGGGCGCGCGCTGGGCGAGCGATACGCGCGAGTGGAGAGAGAG 791
Qy 240 GlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla 259
Db 792 GGGAGCGCGGACCCGAGACCGCTGGACGAGTGAAGAGAGAGCGGGGGGGGGCGCA 851
Qy 260 LysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGln 276
Db 852 CAACTGAAAGAAACAGCCAGAAATACGCTTA---GCCGAGGCTTCAAG 899
```

```
RESULT 9
Bg774871 817 bp mRNA linear EST 15-MAY-2001
LOCUS 602649975F1 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4760827 5',
DEFINITION mRNA sequence.
ACCESSION Bg774871
VERSION Bg774871.1 GI:14045188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRCM1612 row: e column: 20
High quality sequence stop: 813.
Location/Qualifiers
1..817
```

```
FEATURES
Source
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
```

```
/note="Organ: prostate; Vector: pOMB7; Site_1: xhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC library."
```

```
BASE COUNT 148 a 242 c 324 g 103 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1,62e-114	Length:	817
Score:	1182.50	Matches:	251
Percent Similarity:	96.93%	Conservative:	2
Best Local Similarity:	96.17%	Mismatches:	6
Query Match:	74.42%	Indels:	4
DB:	12	Gaps:	0

US-09-827-854-16 (1-317) x Bg774871 (1-817)

```
Qy 1 MetLysValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 40 ATGAAGTTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 99
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 100 GAGCAAGCGGTGGAGACAGAGCGGAGCGGAGCTCGGACAGCAGCAGCAGCAGCAGCAGCAG 159
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGluGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 160 GGGCAGCGCTGGGAACTGGACACTGGGTCTTTGGGATTAACCTGGCTGGGGTGGAGCA 219
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 220 CTGTCTGACAGAGTGCAGAGAGAGCTGTCTGAGTCCAGTACCCAGAGAGAGAGAGAG 279
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db 280 CTGATGAGACGACCATGAAGAGGTTGAAGGCTTACAAATCCGAACTGAGAGAACACTG 339
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 340 ACCCGGATGGCGAGGAGGAGCGGGCGAGCTGTCCAAAGAGCTGAGCGGGCGGAGGCC 399
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 400 CGGCTGGCGCGGACATGGAGAGACGTGTGGCGCGCTGTGTGACATGCCGGCGAGGTG 459
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 460 CAGGCCATGCTCGGCCACAGACACCGAGAGAGTGGCGGTCCGCTCCGCCACCTGGCGCC 519
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 520 AAGCTGGTAAAGCGGCTCTCCGCGATGGCGATGACTCCAGAAAGCGCTGGGAGGTGAC 579
Qy 181 GlnAlaGlyAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 580 CAGGCGGGGCGCGGAGGAGGCGCGGAGCGGCGCTGAGGCCATCCGAGGCGGCGGGG 639
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 640 CCCCTGGTGAAGAGAGCGCGCTGCGGCTGACACTGTGGGTCTCCCTGGCGGCGAG-CCG 698
Qy 221 LeuGlnGluArgAlaGlnAla--TrrPglGluArgLeuArgAlaArgMetGluGlnMetG 240
Db 699 CTACAGAGAGCGGAGCGGCTTGGGCGAGCGGATGGCCCGCGGATGGAGAGATGG 758
Qy 240 LysSerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla 259
Db 759 GCCACCGGAGCG--GAAACGCTGGAGAGAGTGAAGAGACAGGTGGCGGAGGTGGCGGCC 815
```

RESULT 10
 B1600563
 LOCUS 811 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603244936P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
 mRNA sequence.
 ACCESSION B1600563
 VERSION B1600563.1 GI:15493502
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkevics, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1AM11726 row: g column: 10
 High quality sequence stop: 783.
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5287329"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgaag
); Oligo-dT primed using primer 5'-TTTTTATTTTATTTTATTTT-3',
 size-selected for average insert size 2.3 kb and
 normalized to 10^6 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 148 a 245 c 316 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,31e-114 Length: 811
 Score: 1181.00 Matches: 240
 Percent Similarity: 97.97% Conservative: 1
 Best Local Similarity: 97.56% Mismatches: 4
 Query Match: 74.32% Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-16 (1-317) x B1600563 (1-811)
 Oy 1 MetlyValleuTPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 74 ATGAAGGTTCTGTGGTGGCTGGTCTGTCATCTCTGGCAGATGCCAGGCCAAGCTG 133
 Oy 21 GUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpIns 40
 Db 134 GAGCAAGGCGGTGGAGACAGAGCCGAGCTGCCACCAAGCAGACGAGCGAGAGC 193
 Oy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 Db 194 GGGCAGCGCGTGGAGACTGGCACTGGCTTTGGGATTAAGTCCGCTGGGTGACAGCA 253
 Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 254 CTGTCTGAGCAGGTGTCAGAGAGCTGCTCAGCTCCAGAGTACCCAGGAAGTGGAGCGC 313

Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGln 100
 Db 314 CTGATGAGCAGAGACCATGAAAGGCTTCAATGCGAAGTGGAGAACCACTG 373
 Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db 374 ACCCGGTGGCGGAGAGACAGCGGGCAGCGGTCTCCAGAGAGTCCAGCGCGGAGGCC 433
 Oy 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 434 CGGCTGGGCGCGGACATGAGAGACGTGTGCGGCCCTGTGTGCATACCGCGGAGGCTG 493
 Oy 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 494 CAGCGATGCTGGCCCAAGACACCGAGAGAGTGGCGGGTGGCTTGCCTCCACCTGGCC 553
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 554 AAGCTGCGTAAAGCGGCTCTCCGCGATGCGATGACAGAAAGCGCTGGCACTGATAC 613
 Oy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-Gl 200
 Db 614 CAGCGCGGGGCGCGGAGGCGCGCAGCGCGGCTCAGCGCCATCCGAGAGGCTGGGG 673
 Oy 200 YProLeuValGluGlnArgValArgValAlaAlaThrValGlySerLeuAlaGlyGlnPr 220
 Db 674 CCCCTGTGGTGAACAGGGCCGCGTGGCGGCCCACTGTGGGCTCTCGCCAGGACACC 733
 Oy 220 OLeuGlnGluArgAlaGlnAlaTyrPglyLysArgLeuArgAlaArgMetGluGluMetGl 240
 Db 734 GCTACAGAGAGCGGGCCCAAGCGCTGGGCGAGCGCTGGCGCGGCGAGTGAAGCAATGG 793
 Oy 240 YSerArgThrArgAsp 245
 Db 794 CAGCCGGAACCGCGAA 809
 RESULT 11
 B6829472
 LOCUS 845 bp mRNA linear EST 22-MAY-2001
 DEFINITION 60276376BFP1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
 mRNA sequence.
 ACCESSION B6829472
 VERSION B6829472.1 GI:14177059
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 845)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM1790 row: g column: 17
 High quality sequence stop: 829.
 Location/Qualifiers
 1..845
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4899112"
 /clone_lib="NIH_MGC_42"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 253 CTGTCTGACAGGTGACAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAATGAGGGCG 312
QY 81 LeuMetAspGIuThrMetIysGIuLeuIysAlaTyrIysSerGIuLeuGIuGIuLeu 100
313 CTGATGAGACGAGACCATGAGAGAGTGAAGGCTTCAAAATCGAAGCTGGAGAACACATG 372
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerIysGIuLeuGIuAlaGIuAla 120
|||||
Db 373 ACCCGGAGTGGCGAGAGAGCGGGGACGGCTGCTCAAGAGACTGACGGCGGAGGCGC 432
QY 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTyrArgGIuIysGIuVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGTGTGCGGCCCGCTGTGCTACATCCGGCGGCAAGGTG 492
QY 141 GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
493 CAGGCAATGCTGGCGCAGAGCACCGAGGAGCTGGCGCTGCGCTCCACCTGCGCG 552
QY 161 LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGIuIysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTAAGCGGCTCTCCGCGCATGCGCATGACATGACAGAACCGCTGGCACTGTGAC 612
QY 181 GIuAlaGIuAlaArgGIuGIuValGIuGIuArgGIuIysLeuSerAlaIleArgGIuArgLeuGIy 200
613 CAGGGCGGGGGCGCGGAGGGCGCCAGCGCGGTCTCAGCGCATCCGCGAGCGCTGAGG 672
QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuInPro 220
Db 673 CCCCTGGTGGAGACAGAGGGCGCGTGGCGCGGCGGCTGTGCGCTGCGCGAGCGCGG 732
QY 221 LeuGIuGIuArgAlaGIuAlaThrPrlGIuGIuArgLeuArgAlaArgMetGIuGIuMet 239
Db 733 CTACAGGAGCGGGCGGCGGCTGGGGCGAGCGGTG-CGCGCGGATGAGAGATG 788
RESULT 13
BI551475 919 bp mRNA linear EST_05-SEP-2001
LOCUS 603194314P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION BI551475
VERSION BI551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 919)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs@remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
Location/Qualifiers
FEATURES
SOURCE 1.
1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 4,58e-110 Length: 919
Score: 1141.00 Matches: 246
Percent Similarity: 93.26% Conservative: 3
Best Local Similarity: 92.13% Mismatches: 17
Query Match: 71.81% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-16 (1-317) x BI551475 (1-919)
QY 1 MetIysValLeuThrPrlAlaLeuLeuValThrPheLeuAlaGIyCysGIuAlaIysVal 20
|||||
Db 73 ATGAAGTTCTGTGGGCTGCGCTTCTGCTGATTCCTGGAGATGCCAGGCCAAAGGTG 132
QY 21 GIuGIuAlaValGIuThrGIuProGIuProGIuLeuArgGIuGIuThrGIuThrPrlInser 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGGCTGGCGGACAGACAGAGTGGAGAGAC 192
QY 41 GIyGIuArgTPrpGIuLeuAlaLeuGIuArgPheTPrpAspTyrLeuArgTPrpAlaGIuThr 60
|||||
Db 193 GGCCAGCGCTGGAGACTGCGCATGTGGTGTGGATTTACTCTCGCTGGGTGACAGACA 252
QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 253 CTGTCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCAGGTCCAGGATCCAGGAATGTAGAGCGG 312
QY 81 LeuMetAspGIuThrMetIysGIuLeuIysAlaTyrIysSerGIuLeuGIuGIuLeu 100
|||||
Db 313 CTGATGAGACGAGACCATGAGAGAGTGAAGGCTTCAAAATCGAAGCTGGAGAACACTG 372
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerIysGIuLeuGIuAlaGIuAla 120
|||||
Db 373 ACCCGGAGTGGCGAGAGAGCGGGGACGGCTGCTCAAGAGACTGACGGCGGCGAGGCGC 432
QY 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGIuTyrArgGIuIysGIuVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGTGTGCGGCCCGCTGTGCTACATCCGCGGCGAGGTG 492
QY 141 GIuAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
493 CAGGCAATGCTGGCGCAGAGCACCGAGGAGCTGGGGCGGCTGCGCTCCACCTGCGCG 552
QY 161 LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGIuIysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTAAGCGGCTCTCCGCGCATGCGCATGACATGACAGAACCGCTGGCACTGTGAC 612
QY 181 GIuAlaGIuAlaArgGIuGIuValGIuGIuArgGIuIysLeuSerAlaIleArgGIuArgLeuGIy 200
613 CAGGCGGGGGCGGAGTGGCGGCGGCGGCTGACAGCGCATCCGCGAGCGCTGAGG 672
QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuInPro 220
Db 673 CCCCTGGTGGAGACAGAGGGCGCGTGGCGCGGCTGTGCGCTGCGCGAGCGGCGCCT 732
QY 221 LeuGIuGIuArgAlaGIuAlaThrPrlGIuGIuArgLeuArgAlaArgMetGIuGIuMetGI 240
Db 733 A--CAGGAGCGGGGCGGAGCGCTGGGGCGAGCGGTGCGCGCGATGAGAGAAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGIuValIysGIuGIuValAlaGIuValArgAlaIy 260
Db 791 CAGCGGAGCCGGGAGCG--CTGGACGAGTGAAGAGCACGTGGGCGGAGGG-TGGGCCAA 847

QY 260 sleuglugluglnalagln 266
 |||||||::|||
 Db 848 GCTGAGAGACCAAGAG 866

RESULT 14
 BM042228

LOCUS BM042228 757 bp mRNA linear EST 07-NCV-2001
 DEFINITION 603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
 mRNA sequence.

ACCESSION BM042228
 VERSION BM042228.1 GI:16771495

KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 757)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTP/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2021 row: k column: 21
 High quality sequence stop: 757.

FEATURES

source

1..757
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557004"
 /clone_1ib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."

BASE COUNT 134 a 229 c 236 g 98 t

ORIGIN

Alignment Scores:
 Pred. No.: 3.95e-109 Length: 757
 Score: 1131.00 Matches: 234
 Percent Similarity: 98.738 Conservative: 0
 Best Local Similarity: 98.738 Mismatches: 2
 Query Match: 71.188 Indels: 2
 Gaps: 0

US-09-827-854-16 (1-317) x BM042228 (1-757)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||::|||
 Db 50 ATGAAAGTTCTGTGGCTCGTGTGTCACATTCCTGGCAGAGATGCCAGGCCAAGTG 109

QY 21 GlnGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnIns 40
 |||||||::|||
 Db 110 GAGCAAGCGGTGAGACAGAGCCGAGAGCCGAGCTCGCCAGAGACGAGAGCGAGAG 169

QY 41 GlyGlnArgTrpGlnLeuLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||::|||
 Db 170 GGCACAGCGTGGAACTGGCAGCTGGCTCTTTGGGATTACTGGCGCTGGGTCAGACA 229

QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||||::|||
 Db 230 CTGTCTGAGCAGCGTGCAGAGAGCGTCTCAGCTCCAGCTCCAGTCCAGCAAGTACGAGCGCG 289

QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
 |||||||::|||
 Db 290 CTGATGAGCAGACCATGAAAGAGTTGAAGCCTTACAAATCGCAACTGGAGAACTAG 349

QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnGlnAlaAlaGlnAla 120
 |||||||::|||
 Db 350 ACCCGGTTGGCGGAGAGAGACCGCGGACGCTGTCCAAAGAGCTGCAGCGCGGAGCGCC 409

QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 |||||||::|||
 Db 410 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGCTGAGTACCGCGCGGAGTG 469

QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||::|||
 Db 470 CAGGCCATGCTCGCCAGAGACCGAGAGAGCTCGGGGCTCGCTCCCTCCAGCTGCGCC 529

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||||::|||
 Db 530 AACCTGCTGAAGCGGCTCTCTCCGCGATGCCGATGACCTGCAGAGCGCTGGCAGTGTAC 589

QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
 |||||||::|||
 Db 590 CAGCGCGGGCGCCGACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649

QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 |||||||::|||
 Db 650 CCCCTGCTGGACAGAGGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 709

RESULT 15
 BG706129

LOCUS BG706129 907 bp mRNA linear EST 07-MAY-2001
 DEFINITION 602669093F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
 mRNA sequence.

ACCESSION BG706129
 VERSION BG706129.1 GI:13981169

KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 907)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM10669 row: 1 column: 23
 High quality sequence stop: 837.

FEATURES

source

1..907
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4792030"
 /clone_1ib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified

pbuescript K5+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others

ALIGNMENT SCORES:

Pred. No.:	5.14e-108	Length:	907
Score:	1121.50	Matches:	243
Percent Similarity:	94.23%	Conservative:	2
Best Local Similarity:	93.46%	Mismatches:	11
Query Match:	70.58%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-16 (1-317) x BG706129 (1-907)

```
OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 75 ATGAAGGTTCTGTGGCTGCTGTTGCTGCATTCTGGCAGATGCCAGGCAAGGTG 134
OY 21 GlnGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 135 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCCAGACAGACCGAGTGGCAGAC 194
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 195 GGCCAGCGCTGGAGACTGGACATGGCTGCTTTGGATTACCTGCGCTGGGTGCAGACA 254
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 255 CTGTGTGAGCAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGCG 314
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 315 CTGATGGAGACAGACCATGATGAGAGTTGAAAGCCTTCAAAATCGAATCGAGACACACTG 374
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
DB 375 ACCCGGTGGCGGAGAGCGCGGACGCGCTGTCCAAAGAAAGTGCAGGCGGCGCAGGC 434
OY 120 ArgGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 435 CCGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCCCTGTGTGACGTACCGCGCGAGGT 494
OY 140 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 495 GCGAGCCATGCTCGCGCCAGAGACCGAGAGAGTGGCGGTGCGCTGCGCTCCACCTGCG 554
OY 160 GlysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 555 CAAGCTGCGTAAAGCGGCTCTCCGGATGCCATGACCTGCAGAGCGCGCTGGCAGTGA 614
OY 180 rGlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-G 200
DB 615 CCAGGCCGCGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCCATCCGAGCGCGCTGGG 674
OY 200 LysProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnP 220
DB 675 GCCCTTGTGTGAACAGGCGCGCGTGGCGGCGCCACTGTGGGCTCCTGGAGCGCAGAC 734
OY 220 rLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetG 240
DB 735 CGCTACAGAGAGCGCGCAGGCTG--GGGCGAGCGGTGCGCGCGGATGAGAGAGATGG 791
OY 240 LysArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArg 258
DB 792 GCAG-CGAGCCCGG--ACGCTGACGAGTG- AAGGAGCAGTGGGGAAGTGTGCGC 842
```

Search completed: March 14, 2003, 20:14:03
Job time : 1293.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-17
Perfect score: 1589
Sequence: 1 MKVLMALLVFLAGCOAKV.....VEKVAAGVTSAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/sgn2.1/USPRO/spool/US09827854/runat_11032003.101610.27486/app.query.fasta.1.3576
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 @cgn.1.13008 @runat.11032003.101610.27486 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -TARGETQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1582	99.6	1156	6	BD004277 Apo E hum
2	1570	98.8	1110	6	E00359 E00359 cDNA coding
3	1570	98.8	1110	6	E00823 DNA sequence
4	1570	98.8	1147	6	AX302545 Sequence
5	1570	98.8	1156	6	BD004278 Apo E hum
6	1570	98.8	1156	9	BD004278 Apo E hum
7	1570	98.8	1186	9	K00396 Homo sapien
8	1563	98.4	1110	6	BC003557 Homo sapi
9	1561	98.2	1157	6	E08423 DNA coding
10	1561	98.2	1157	6	AX333278 Sequence
11	1561	98.2	1157	6	AX409597 Sequence
12	1561	98.2	1157	6	AX409597 Sequence
13	1558	98.0	1156	6	HMABPOE HMABPOE
14	1436	90.4	1178	6	BD004279 Apo E hum
15	1381.5	86.9	5491	9	MFAPOR MFAPOR
16	1381.5	86.9	41907	6	AF261279 AF261279
17	1381.5	86.9	41907	6	AX358722 Sequence
18	1381.5	86.9	41907	9	AF050154 Homo sapi
19	1364	85.8	5515	9	AC011481 Homo sapi
20	1335	84.0	5413	9	AC011481 Homo sapi
21	1252	78.8	4762	9	AF261280 Pan trogl
22	1251	78.7	208239	2	BAAPOE BAAPOE
23	1243	78.2	1138	4	AC021988 Sequence
24	1155	72.7	1060	4	AF303830 Tupia gl
25	1148.5	72.3	965	6	RRABPOLP RRABPOLP
26	1148.5	72.3	1108	4	AX384545 Sequence
27	1148.5	72.3	5617	6	AX384541 Sequence
28	1148.5	72.3	6026	6	AX384539 Sequence
29	1136	71.5	718	9	AX384539 Sequence
30	1134	71.4	1045	10	AF200497S3 Pan trogl
31	1134	71.4	1104	10	MUSAPOE MUSAPOE
32	1132	71.2	718	9	BC028816 Mouse apoli
33	1132	71.2	718	9	AF200500S3 Gorilla g
34	1130.5	71.1	1154	4	AF200503S3 Pongo pyg
35	1128	71.0	718	9	AF200506S3 B. taurus mr
36	1128	71.0	1122	4	BRAPOE BRAPOE
37	1123	70.7	1126	6	SSAPOR SSAPOR
38	1123	70.7	1126	6	AR164342 Sequence
39	1079	67.9	951	10	AR205885 Sequence
40	1057	66.5	959	10	S76779 Raboe-apol1
41	1041.5	65.5	1069	10	MUSAPOE MUSAPOE
42	1014	63.8	228698	2	AC127479 Mus muscu
43	1014	63.8	237653	2	AC073760 Mus muscu
44	988	62.2	4856	10	AC073760 Mus muscu
45	987	62.1	4267	6	AR164387 Mus muscu

RESULT 1

ALIGNMENTS

LOCUS	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Apo E humanized mammal.				
ACCESSION	BD004277				
VERSION	BD004277.1 GI:18632238				
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.				
TITLE	Apo E humanized mammal				
JOURNAL	Patent: JP 2001017028-A 1 23-JAN-2001;				
COMMENT	MITSUBISHI CHEMICAL CORP				
OS	Homo sapiens (human)				
PN	JP 2001017028-A/1				
PD	23-JAN-2001				
PF	28-APR-2000 JP 2000128919				
PR					
PI	SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA				
PC	A01K87/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC				
PC	C12N15/09//C07K14/775,				
PC	(C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N15/00,C12R1:91) CC				
FEATURES					
source	FT	key	Location/Qualifiers		
	CDS		Location/Qualifiers		
			1..1156		
			/organism="Homo sapiens"		
			/db_xref="taxon:9606"		
BASE COUNT	208 a	367 c	432 g	149 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	1,59e-84	length:	1156		
Score:	1582.00	Matches:	316		
Percent Similarity:	99.68%	Conservative:	0		
Best Local Similarity:	99.68%	Mismatches:	1		
Query Match:	99.56%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-17 (1-317) x BD004277 (1-1156)					
OY	1	Methylsvalleu	ThrpAlaAlaLeuValThrpheLeuAlaGlyCysGlnAlaLysVal	20	
Db	61	ATGAGGTTCTGTGGGCTGGCTTGTGTCACATCTCGCAGATGCGCAGGCCAAGGTG	120		
OY	21	GIUGlnAlaValGIUthGluProGluProGluLeuArgGlnGlnthGluTrpGlnser	40		
Db	121	GAGCAGGCGGTGGACAGAGCCGAGCCGCGAGCTGCGCCACAGACCGAGTGGCGAGC	180		
OY	41	GIUGlnArgTrpGluLeuAlaLeuGlnIlyArpPheTrpAspTyrLeuArgTrpValGlnThr	60		
Db	181	GGCCAGCGCTGGGAACCTGGCACCTGGGTGGCTTTTGGATTACCTGCGCTGGGTGCAGACA	240		
OY	61	LeuSerGIUGlnValGIUGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80		
Db	241	CTGTCTGACAGAGTCAGAGAGGAGCTGCTCAGACCTCCAGGTCACCCAGAGAACTGAGGCGC	300		
OY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGIUGlnGluGlnLeu	100		
Db	301	CTGATGGACGAGACCATGAAAGAGTTGAAGGCTTCAAAATCGGAACCTGAGAGAACTG	360		
OY	101	ThrProValAlaGlnGluThrArgAlaIleArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
Db	361	ACCCGCGGTGGCGAGAACCGCGGCGACGCTGTCTCAAGAGACCTGCAGGCGCGAGGCC	420		
OY	121	ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal	140		
Db	421	CGGCTGGGCGGAGCATGGAGAGCTGTGGCGGCCCTGTGTGACGTACCGGCGGAGGTG	480		
OY	141	GlnAlaMetLeuAspGlnSerThrGluLeuArgValAlaArgLeuAlaSerHisLeuArg	160		

DB	Accession	Gene	Protein	Length	Species	Source	Comment
Db	481	CAGGCCATCTCTGGCCAGAGCACCGAGAGCTGCGGGTCCGCTCCACCTGGCC		540			
QY	161	LysLeuAtrLysArgLeuLeuAtrgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr		180			
Db	541	AAGCTGCTGAACCGGCTCTCTCCGCGATGCGGATGACCTCAGAAAGTGGCTGGCAGGTAC		600			
QY	181	GlnAlaGlyAlaArgGluGluValAlaGluAtrgIleuSerAlaIleArgGluAtrgLeuGly		200			
Db	601	CAGCGCGGGGGCCCGAGAGGGCCCGAGCGGGCTCTACGCGCATCCGCGAGCGCTGGGG		660			
QY	201	ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro		220			
Db	661	CCCCTGGTGGAACAGAGGGCCGCTGCGGGCCGACCTGTGGCTCCCTGGCCGCGCAGCGG		720			
QY	221	LeuGlnGluAtrgAlaGlnAlaIleAtrpGlyGluAtrgLeuAtrgAlaArgMetGluMetGly		240			
Db	721	CTACAGAGACGGGGCCAGAGCTGTGGGCGAGCGGCTGCGCGCGCGATGTAGAGAGATGGCG		780			
QY	241	SerAtrgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys		260			
Db	781	AACCGGACCCCGGAGCGGCTGTGACAGAGGTGAAGAGACAGTGGCGGAGGTGGCGCGCAAG		840			
QY	261	LeuGluGlnAlaGlnAlaGlnIleAtrgLeuGlnAlaGlnAlaPheGlnAlaAtrgLeuLys		280			
Db	841	CTGGAGAGACAGGCCCGACAGATACGCTGTGACAGCGCGAGGCTTCCAGGCCGCTCAAG		900			
QY	281	SerTrrpPheGluProLeuValGluAspMetGlnAtrgGlnTrrpAlaGlyLeuValGluLys		300			
Db	901	ACCTGGTTCGACGCCCTGTGTGAGACATGACAGCGCGCACTGGGCGGGCTGTGGAGAG		960			
QY	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis		317			
Db	961	GTCAGGCTGCGCTGGGCGACCGCGCCGCTGTGCGCCAGCAGCAATCTAC		1011			
RESULT 2							
LOCUS	E00359		1110 bp	RNA	linear	PAT 29-SEP-1997	
DEFINITION	E00359						
ACCESSION	E00359.1	GI:2168646					
VERSION	JP 1985118189-A/1.						
KEYWORDS	Homo sapiens.						
SOURCE	Homo sapiens.						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	Teraishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.						
TITLE	DNA FRAGMENT						
JOURNAL	Patent: JP 1985118189-A 1 25-JUN-1985;						
COMMENT	MITSUBISHI CHEM IND LTD						
	OS human						
	PN JP 1985118189-A/1						
	PD 25-JUN-1985						
	PF 29-NOV-1983 JP 1983224980						
	PI TERAIISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI						
	KIMURA MASAKO,						
	PI IKEDA YASUOKO						
	PC C12N15/00.C07H21/04//C12P21/00;						
	CC strandedness: Double;						
	CC topology: Linear;						
	CC hypothetical: No;						
	CC anti-sense: No;						
	CC *source: tissue-type=livr;						
	FH key location/Qualifiers						
	FT CDS 15..968						
	FT /product='apolipoprotein E3 precursor' FT						
	FT sig_peptide 15..68						
	FT /product='apolipoprotein E3 signal peptide' FT						
	FT mat_peptide 69..965						
	FT /product='apolipoprotein E3'.						
	FT location/Qualifiers						

source		1. .1110		/organism="Homo sapiens"		/db_xref="taxon:9606"	
BASE COUNT	198 a	353 c	416 g	143 t			
ORIGIN							
Alignment Scores:							
Pred. No.:	7.67e-84	Length:	1110				
Score:	1570.00	Matches:	315				
Percent Similarity:	99.37%	Conservative:	0				
Best Local Similarity:	99.37%	Mismatches:	2				
Query Match:	98.80%	Indels:	0				
DB:	6	Gaps:	0				
US-09-827-854-17 (1-317) x E00359 (1-1110)							
QY	1	MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20				
DB	15	ATGAAGGTTCTGTGGGCTCGTGTGCTGCATTCCTGCGAGATGCCAGGCCAAAGGTG	74				
QY	21	GIuGlnAlaValaGIuThrcGIuProGIuProGIuLeuArgGlnGlnThrcGIuTrpGlnSer	40				
DB	75	GAGCAAGCGGTGGAGACAGAGCCGGAGCCGAGCTGCGCAGAGAGACGAGTGGCAGAC	134				
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60				
DB	135	GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGAGATTACCTGGCGTGGGTGCAGACA	194				
QY	61	LeuSerGIuGlnValaGIuGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80				
DB	195	CTGTCTGACAGGTGCGAGAGAGAGCTGCTCAGCTCCAGAGTCAACCGAGACTGAGGGCG	254				
QY	81	LeuMetAspGIuThrcMetLysGIuLeuLysAlaTyrLysSerGIuLeuGlnGluGlnLeu	100				
DB	255	CTGATGAGACGACCATGAAAGAGATTGAAGGCTTACAAATCCGAACTGGAGAGACACTG	314				
QY	101	ThrProValaIaGIuGluThrcArgAlaArgLeuSerLysGIuLeuGlnAlaIaGlnAla	120				
DB	315	ACCCGCGTGGCGAGAGAGCGGGCGCGCTGTCCAAAGAGCTGCAAGCGCGCGAGCC	374				
QY	121	ArgLeuGlyAlaAspMetGIuAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal	140				
DB	375	CGGCTGGGGCGGACATGAGAGAGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG	434				
QY	141	GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160				
DB	435	CAGGCGATGCTCGGGCCAGAGACACCGAGAGACTCGGGTGGCTGCCCTCCACCTGGCC	494				
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180				
DB	495	AAGCTGCGTAAGCGGGCTCTCCGCGATGGCGATGACCTGCAGAGCGCTGGCAGGTAC	554				
QY	181	GlnAlaGlyAlaArgGlnGlyAlaGlnGlyLeuSerAlaIleArgGlnArgLeuGly	200				
DB	555	CAGGCGGGGGCGCGAGGGCGCGCGAGCGGGCTCAGCGCCATCCGCGAGCGCGCTGAGG	614				
QY	201	ProLeuValaGIuGlnGlyArgValaArgAlaIaThrValGlySerLeuAlaGlnPro	220				
DB	615	CCCCTGTGTGAACAGGGCGCGCTGCGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTG	674				
QY	221	LeuGlnGluArgAlaGlnAlaIaTrpGlyGluArgAlaArgMetGlnGlnMetGly	240				
DB	675	CTACAGAGAGCGGGCGCGAGCGCTGGGGCGAGCGGTGGCGCGGATGAGAGATGGGC	734				
QY	241	SerArgThrArgAspArgLeuAspGlnValLysGIuGlnValaIaGlnValaIaArgAlaIys	260				
DB	735	ACCCGAGCCCGGACCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	794				
QY	261	LeuGlnGluGlnAlaGlnGlnIleArgGlnGlnAlaGlnAlaPheGlnAlaIaArgLeuLys	280				
DB	795	CTGGAG	854				
QY	281	SerTrpPheGluProLeuValaGlnAspMetGlnArgGlnTrpAlaGlyLeuValaGlnLys	300				

DB	855	AGCTGTTTCAGACCCCTGTGTGAGAGACATCAGCCCATGTGGGCTGTGTGAGAGAG	914		
QY	301	ValGlnAlaAlaValaGIyThrcSerAlaAlaIaProValProSerAspHis 317			
DB	915	GTGCAGGCTGGCGGTGGGACAGAGCGCGCCCTGTGTGCCAGGACATCATC	965		
RESULT 3					
E00823	E00823	1110 bp	DNA	linear	
LOCUS					
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal peptide.				
ACCESSION	E00823				
VERSION	E00823.1	GI:2169084			
KEYWORDS	JP 1986096997-A/1.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1110)				
AUTHORS	Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.				
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN				
JOURNAL	Patent: JP 1986096997-A 15-MAY-1986;				
COMMENT	MITSUBISHI CHEM IND LTD				
OS	Human (Homo sapiens)				
PN	JP 1986096997-A/1				
PD	15-MAY-1986				
PF	16-OCT-1984 JP 1984216987				
PI	TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO				
PC	C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12N1:19), (C12N15/00,				
PC	C12N1:19);				
CC	strandedness: Double;				
CC	topology: Linear;				
CC	hypothetical: No;				
CC	anti-sense: No;				
CC	*source: tissue=livr;				
FT	key				
FT	Location/Qualifiers				
FT	3'UTR 1..14				
FT	sig.peptide 15..68				
FT	product=human apolipoprotein E signal				
FT	peptide				
FT	69..968				
FT	CDS				
FT	product=human apolipoprotein E' FT 3'UTR				
FEATURES					
source	1. .1110				
	Location/Qualifiers				
	969..1110.				
	/organism="unidentified"				
	/db_xref="taxon:32644"				
BASE COUNT					
198 a	353 c	416 g	143 t		
ORIGIN					
Alignment Scores:					
Pred. No.:	7.67e-84	Length:	1110		
Score:	1570.00	Matches:	315		
Percent Similarity:	99.37%	Conservative:	0		
Best Local Similarity:	99.37%	Mismatches:	2		
Query Match:	98.80%	Indels:	0		
DB:	6	Gaps:	0		
US-09-827-854-17 (1-317) x E00823 (1-1110)					
QY	1	MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20		
DB	15	ATGAAGGTTCTGTGGGCTCGTGTGCTGCATTCCTGCGAGATGCCAGGCCAAAGGTG	74		
QY	21	GIuGlnAlaValaGIuThrcGIuProGIuProGIuLeuArgGlnGlnThrcGIuTrpGlnSer	40		
DB	75	GAGCAAGCGGTGGAGACAGAGCCGGAGCCGAGCTGCGCAGAGAGACGAGTGGCAGAC	134		
QY	41	GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr	60		
DB	135	GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGAGATTACCTGGCGTGGGTGCAGACA	194		

DB 855 ACCTGGTTCAGACCCCTGTGTGAAGACAGCAGCCGCACTGGCGGCTGTGTGAAGAG 914

QY 301 ValGlnAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317

DB 915 GTGCAAGCTGCGGTGGGACAGAGCCGCCCTGTGTGCCAGAGACATCAC 965

RESULT 3

LOCUS E00823

DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.

ACCESSION E00823

VERSION E00823.1 GI:2169084

KEYWORDS JP 1986096997-A/1.

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1110)

AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.

TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN

JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;

COMMENT MITSUBISHI CHEM IND LTD

OS Human (Homo sapiens)

PN JP 1986096997-A/1

PD 15-OCT-1984 JP 1984216987

PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO

PC C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12N1:19),

PC (C12N15/00,

PC C12R1:19);

CC strandedness: Double;

CC topology: linear;

CC hypothetical: No;

CC anti-sense: No;

CC *source: tssuse-type=Liver;

key Location/Qualifiers

FT 3'UTR 1..14

FT sig-peptide 15..68

FT peptide' /product='human apolipoprotein E signal FT

FT CDS 69..968

FT /product='human apolipoprotein E' FT 3'UTR

FEATURES

source 969..1110.

location/Qualifiers

1..1110

/organism='unidentified'

/db_xref='taxon:32644'

BASE COUNT 198 a 353 c 416 g 143 t

ORIGIN

Alignment Scores:

Pred. No.: 7.67e-84 Length: 1110

Score: 1570.00 Matches: 315

Percent Similarity: 99.37% Conservative: 0

Best Local Similarity: 99.37% Mismatches: 2

Query Match: 98.80% Indels: 0

DB: 6 Gaps: 0

US-09-827-854-17 (1-317) x E00823 (1-1110)

QY 1 MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20

DB 15 ATGAAGGTTCTGTGGGCTCGTGTGCTGCATTCCTGCGAGATGCCAGGCCCAAGGTG 74

QY 21 GIuGlnAlaValaGIuThrcGIuProGIuProGIuLeuArgGlnGlnThrcGIuTrpGlnSer 40

DB 75 GAGCAAGCGGTGGAGACAGCCGGAGCCGAGCTGCGCAGAGAGACGAGTGGCAGAGAC 134

QY 41 GIyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 135 GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGAGATTACCTGGCGTGGGTGCAGACA 194

BASE COUNT	210 a	365 c	425 g	147 t	ORIGIN
Alignment Scores:					
Pred. No.:	7.94e-84			1147	
Score:	1570.00			315	
Percent Similarity:	99.37%			0	
Best Local Similarity:	99.37%			2	
Query Match:	98.80%			Indels:	
DB:	6			gaps:	0
US-09-827-854-17 (1-317) x AX302545 (1-1147)					
Oy	1	MethylValIleuTPAlaAlaLeuLeuValThrpheLeuAlaGlyCysGlnAlaIalysVal	20		
Db	46	ATGAAGGTTCTGTGGCGCTTCTGCTGCATTCTCGAGAGATCCAGGCCAAGCTG	105		
Oy	21	GlulGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40		
Db	106	GAGCAAGCGGTGGAGACAGACCCGAGCCAGCTGCGCCAGCAGACCGAGTGGCAAGC	165		
Oy	41	GlyGlnArgTrpGluLeuAlaAlaLeuGlyArgPheTrpAspPyrLeuArgTrpValGlnThr	60		
Db	166	GCCAGCGCTGGAACTGGCACTGGTGGCTTTGGGATTAACCTGGCCCTGGGTGCACACA	225		
Oy	61	LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80		
Db	226	CTGTCTGAGCAGGTGGAGAGAGACTGCTCAGCTCCAGGTACACCAGAACTGAGGGCG	285		
Oy	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu	100		
Db	286	CTGAAGGACGAGACCATTGAAGAGATTGAAGCCCTCAAAATCGGAACGGAGAAACACTG	345		
Oy	101	ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120		
Db	346	ACCCTGGCGGGAGAGAACCCGGGGCAGCGTGTCCAAAGACGTGCAGCGCGCCAGGCC	405		
Oy	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal	140		
Db	406	CGCGTGGCGCGGACATGTGAAGAGCTGTGGCGCCCTGTGGTACGTACCCCGGAGAGTG	465		
Oy	141	GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160		
Db	466	CAGGCCATTGCTCGGCCAGACACCAGAGAGCTGGCGGGTGGCTTCCACTTCGCG	525		
Oy	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180		
Db	526	AAGCTGCTAAGCGGCTCTCCGCGATGCCGATACCTGCAGAAAGCCCTGGCAGTGTAC	585		
Oy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200		
Db	586	CAGCGCGGGGCCCGCAGAGGGCGCGCAGCGGCTTCAGCGCCATCCCGAGCCCTGGGG	645		
Oy	201	ProLeuValGlnGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlnPro	220		
Db	646	CCCCGTGGTGAACAGGGCGCGGTGCGGGCCGCCACTGTGGGTCTCCGTGGCGGCCAGCG	705		
Oy	221	LeuGlnGlnArgAlaGlnAlaIleArgLysGlyLysArgLeuArgAlaArgMetGlnGluMetGly	240		
Db	706	CTACAGAGCGCGGCCCGAGGCTGTGGGGAGAGCGCTGCGCCGCGCGATGAGAGATGGGC	765		
Oy	241	SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys	260		
Db	766	AGCCGGACCCCGACCGCTGTGACGAGGTGAAGAGCAGGTGGGGAGGTGGCGCCAAG	825		
Oy	261	LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280		
Db	826	CTGGAGGAGCAGGCCACAGACATACGCTGTGAGGCCAGAGGCTTCCAGGCCCGGCTCAAG	885		
Oy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys	300		
Db	886	AGCTGGTTCGAGCCCTGTGTGGAAGACATGACAGCGCAAGTGTGGCGGGCTGTGTGAGAACG	945		

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 946 GTGACAGCTCGCTGGGACACAGCGCCCTGTGTGCCAGACAAATCAG 996

RESULT 5
 BD004278

LOCUS BD004278 1156 bp DNA linear PAT 31-JAN-2002
 DEFINITION Apo E humanized mammal
 ACCESSION BD004278
 VERSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1156)
 Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
 Apo E humanized mammal
 Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PF 23-JAN-2001
 PD 28-APR-2000 JP 2000128919
 PR
 PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA
 A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
 C12N15/09, C07K14/775,
 PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

FEATURES
 source 1..1156 Location/Qualifiers
 FT CDS (61)..(1011).
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t

ORIGIN

Alignment Scores:
 Pred. No.: 8.01e-84 Length: 1156
 Score: 1570.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.80% Indels: 0
 Gaps: 0

US-09-827-854-17 (1-317) x BD004278 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGTCTGTGGGCTGCGTTCGTCGACATTCCTGCGAGATGCCAGCGCAAGGTG 120

QY 21 GlnGlnAlaValGlnThrGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGCGGAGCCGAGACTGCGCAGAGAGACGAGTGGCAGAC 180

QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 181 GGCACAGCTGGGAACTGGACACTGGGTGCTTTGGATTACCTGCGGTGGGCGAGACA 240

QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 241 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGAGACTGAGGGCG 300

QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 301 CTGATGAGAGACCATGTAAGAGAGTTGAAGCGCTTACAATTCGAAGTGAAGAGACAATG 360

QY 101 ThrProValAlaGlnGlnThrArgAlaAlaGlnSerLysGlnLeuGlnAlaAlaGlnAla 120
 Db 361 ACCCCGCTGGCGAGAGACGCGGGCGACGCTGTCCAAAGAGACTGACGCGCGCGCGCC 420

QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 Db 421 CGGCTGGGGCGGACATGAGAGACGCTGGGGCGGCTGTGCATGACCGCGGCGAGGTG 480

QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCGCATGCTCGGACAGACAGACGAGCTCGGGTGGCTCCCTCCCTCCACCTGCGC 540

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 541 AAGCTCGTAAAGCGGCTCTCGCGGATGCCGATGACCTGCAGAAAGCCCTGCGAGTGC 600

QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 Db 601 CAGCGCGGGCGCGCAGAGGGCGCGGAGCGGCTCAGCGCCATCGCGCAGCGCCCTGGGG 660

QY 201 ProLeuValGlnGlnGlnArgValAlaGlnAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCCCTGCTGGAACAGGCGCGCGTGGCGCGCCACTGTGTGGCTCCCTGCGCGCGCAGCGG 720

QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 721 CTACAGAGAGCGGCGCAGGCTGGGCGAGCGGCTGCGCGCGGATGGAGATGGCG 780

QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 Db 781 AGCGGACCGCGGACCGCGCTGGAGCGAGTGAAGACAGAGTGGCGGAGGTGGCGCGCAAG 840

QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGGAGAGACAGCGCCAGCAGATACCTGAGGCGCGAGCGCTCCAGCGCCCTCCACAG 900

QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
 Db 901 AGCTGTTGAGCGCTCTGTGTAAGACATGACAGCGGCTGAGGCGGCTGTGTGAGAAAG 960

QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGACAGCTCGCTGGGACACAGCGCCCTGTGTGCCAGACAAATCAG 1011

RESULT 6
 HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 LOCUS HUMAPOE3
 DEFINITION Homo sapiens preapoliipoprotein E (APOE) mRNA, complete cds.
 ACCESSION K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apoliipoprotein; apoliipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 355 to 1156)
 Breslow, J.L., McPherson, J., Nussbaum, A.L., Williams, H.W.,
 Lofquist-Kahl, F., Karathanasis, S.K. and Zanits, V.I.
 Identification and DNA sequence of a human apoliipoprotein E cDNA clone

JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404

REFERENCE 2 (bases 250 to 777)
 Wallis, S.C., Rogne, S., Gill, L., Markham, A., Edge, M., Woods, D.,
 Williamson, R. and Humphries, S.
 The isolation of cDNA clones for human apoliipoprotein E and the
 detection of apoE RNA in hepatic and extra-hepatic tissues
 EMBO J. 2 (12), 2369-2373 (1983)

JOURNAL 84131952
 MEDLINE 6199196
 PUBMED

REFERENCE 3 (bases 1 to 1156)
 Zanits, V.I., McPherson, J., Goldberger, G., Karathanasis, S.K. and
 Breslow, J.L.
 Synthesis, intracellular processing, and signal peptide of human
 apoliipoprotein E

Db 315 ACCCCGGTGGCGAGAGACGCCGGCCACGGCTGTCCAAAGAGCTGCAGCGCGCCAGGCC 374
QY 121 ATGLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 375 CGGGTGGGGCGGACATGGAGAGCTGTGGCCGCCCTGTGGTCACTACCGCGCCAGAGTG 434
QY 141 GlnAlaMetLeuAspGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGGCGCATGCTGGGCAAGCAGCAGCAGCGCGGCTGGCTGGCTCCACTGGCGC 494
QY 161 LysLeuArgValArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 495 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGTGCAGAAACGCCGTGGCACTGTAC 554
QY 181 GlnAlaGlyAlaArgGlnGlyValaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 555 CAGGCGGGGGCGCGGAGGGGGCGGAGCGCGCTGCACGCCCATCCGCGAGCGCGCTGGGG 614
QY 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 615 CCCCTGGTGGAAACAGCGCGCGGTGGCGCGCACTGTGGCTCCCTGGCGCGCCAGCGC 674
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 675 CTACAGGAGCGGGCCAGCGCTGGGGCGAGCGCGCTGGCGCGATGGAGAGATGGGC 734
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValaGlnValArgAlaLys 260
Db 735 AGCCGAGACCGCGACCGCTGGAGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCCAAG 794
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 795 CTGGAGGAGCAGCGCCACAGATACGCTGCAGCGCGAGCGCTTCAGCGCCCTCAAG 854
QY 281 SerTrrpPheGluProLeuValGlnAspMetGlnArgGlnTrrpAlaGlyLeuValGlnLys 300
Db 855 AGCTGTGTCGAGCCCTGCTGGTGAAGACATGCAGCGCCAGTGGCGCGCTGGTGGAGAAG 914
QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 915 GTGCAGGCTGCCGTGGGCAACGCGCCGCCCTGTGTGCCAGGACAAATCAC 965
RESULT 9
AX333278 1157 bp DNA linear PAT 09-JAN-2002
AX333278
LOCUS Sequence 3787 from Patent W00194629.
DEFINITION AX333278
ACCESSION AX333278.1 GI:18123912
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R., and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
FEATURES
source location/Qualifiers (US)
1. 1157
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 2.69e-83 Length: 1157
Score: 1561.00 Matches: 313
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.24% Indels: 0

DB: 6 Gaps: 0
US-09-827-854-17 (1-317) x AX333278 (1-1157)
QY 1 MetLysValLeuTrrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCGTGTGGGCTGGCTGGTGTGCATCTCCTGGCAGATGTGCAGGCCAAGGTG 121
QY 21 GlnGlnAlaValaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrrpInsr 40
Db 122 GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCACAGCAACCGATGGCGAGAGC 181
QY 41 GlyGlnArgTrrpGlnLeuAlaLeuGlnValArgPheTrrpAspTrrpLeuArgTrrpValGlnThr 60
Db 182 GGCCAGCCCTGGGAACCTGGCACTGGGTGGCTTTTGGATTACCTGGCTGGGTGGAGACA 241
QY 61 LeuSerGlnGlnValaGlnGlnLeuLeuSerSerGlnValaThrGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCACCAAGAACTGAGGGCG 301
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTrrpLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGACAGAGACCATGAGAGAGTTGAAGGCTTCAATGTGGAACTGGAGAACACTGG 361
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGAGAACCGCGGCAAGCTGTCCAAAGAGCTGCACAGCGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTrrpArgGlyVal 140
Db 422 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGGAGTACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuAspGlnSerThrGlnLeuLeuArgValaGlnAlaAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTGGCGCAGAGCAGCAGAGAGCTGGCGGTGGCTCCCTCCACTGGCGC 541
QY 161 LysLeuArgValArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 542 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGTGCAGAGAGCGCTGGCACTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGlyValaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 602 CAGCGCGGGCGCGGAGCGCGCGAGCGCGGCTGCACGCCCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGAACAGCGCGCGGTGGGGCGCGCACTGTGTGGCTCTCCGTGGCGCGAGCGC 721
QY 221 LeuGlnGlnArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 722 CTACAGGAGCGGGCCAGCGCTGGGGCGAGCGCGCTGGCGCGGATGGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValaGlnAlaValaArgAlaLys 260
Db 782 AGTCGAGACCCCGACCGCTGGTGAAGAGAGAGAGTGGGGAGAGTGGCGCCCAAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGGAGCAGCGCCACAGATACGCTGCAGCGCGAGCGCTTCAGCGCGCTCAAG 901
QY 281 SerTrrpPheGluProLeuValGlnAspMetGlnArgGlnTrrpAlaGlyLeuValGlnLys 300
Db 902 AGCTGTGTCGAGCCCTGCTGGTGAAGACATGTGCAGCGCCAGTGGCGCGCTGGTGGAGAAG 961
QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCCGTGGGCAACGCGCGCCCTGTGTGCCAGGACAAATCAC 1012
RESULT 10
AX409597 1157 bp DNA linear PAT 14-JUN-2002
AX409597
LOCUS Sequence 2244 from Patent W00229103.
DEFINITION AX409597
ACCESSION

[illegible]

Oy	221	LeuglIngluaIAGlaGlAlaIatPrGluYluYrGuLeuAArgMaMetCgUlmwTctly	240
Db	722	CTACAGAGAGCGGGCCCAAGGCTGTGGGGCCAAGGGCGCTGCCGCGGGATGAGAGAATTGGGC	781
Oy	241	SerArqTrArqAsPArqLeuAsPcJuvAlLySgUglInVaLAlaGlVAlArqAlaLyS	260
Db	782	ACTGGGACC CGCGAGCCGCTGGACAGAGTTGAAGAGACAGTGCGGAGGTGCGGCCCAAG	841
Oy	261	LeuglUGlUGlnAlaGlNGlnIlleaRguEuglInalagluAlaIapheGlnAlaArqLeuLyS	280
Db	842	CTGGAGGAGACAGGCCACCATGATACGCTTCAGAGCCGAGGCGCTTCACAGCCGCTCAAG	901
Oy	281	SerTrpPheGluProLeuValGIuAsPmetGlnArqGIntnTPraAgLYleuValGLuLyS	300
Db	902	AGCTGGTTTCGAGCCCCCTGTGTGGAAACATGCAAGCCCCAGTGGCGCGGCTGTGTGAGAAG	961
Oy	301	ValGlnAlaAlaValaIGlyThrSerAlaIaIaProvalProSerAsPaSnHis	317
Db	962	GTCGAGGCTGCCGTGTGGCACACGCGCCGCTGTGTGCCACAGCAATCACC	1012
RESULT	11		
LOCUS	115975	1157 bp	DNA Linear PAT 03-APR-1996
DEFINITION	Sequence 1 from patent US 5472858.		
ACCESSION	115975		
VERSION	115975.1	GI:1250883	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1157)		
AUTHORS	Attie,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.		
TITLE	Production of recombinant proteins in insect larvae		
JOURNAL	Patent: US 5472858-A 1 05-DEC-1995;		
FEATURES	Location/Qualifiers		
Source	1..1157		
BASE COUNT	212 a 370 c 426 g 149 t		
ORIGIN	/organism="unknown"		
Alignment Scores:			
Pred. No.:	2.69e-83	Length:	1157
Score:	1561.00	Matches:	313
Percent Similarity:	98.74%	Conservative:	0
Best local Similarity:	98.74%	Mismatches:	4
Query Match:	98.24%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-17 (1-317) x 115975 (1-1157)			
Oy	1	MettysValleuTriPraIAalaLeuLeuValThrpheLuAlaGlCySGlnAlaLySVal	20
Db	62	ATGAAGATTCTGTGGGCTGCGCTGTGTCATTCCTCGGACAGATGCCAGCCAAAGTG	121
Oy	21	GIuGlnAlaValaGIuThrGIuProGluProGluLeuArqGlngInThrGIuTrpGlnSer	40
Db	122	GAGCAAGCGGTGAGACAGACCAGGCCCAAGCTGCSCCAGACAGCAAGTAGTCGACAGC	181
Oy	41	GIyGlnArqTrpGIuLeuAlaLeuGIyArqrherTpaspPyrrleuArqTrpValGIomThr	60
Db	182	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTTGGGATTTACCTGCGCTGGGTGCAGCA	241
Oy	61	LeuSerGIuGlnValaGIuGIuLeuLeuSerSerGlnValIThrGlngIleuArqAla	80
Db	242	CTGTCTGAGCAGAGTCAGAGGAGAGCTGCTCAAGTCCACCACAAGAACTGAGGCGC	301
Oy	81	LeuMeLaspGIuThrMetLySgluLeuLySaIaTryLySerGIuLeuGIuGlnIleu	100
Db	302	CTGATGAGCAGACCATGAAGAGATTGAAGCCCTTAACAATCGGAACCTGAGAACAACTG	361
Oy	101	ThrpProValaIaGIuGIuThrArqAlaArqLeuSerLySgluLeuGlnAlaIaGlAla	120
Db	362	ACCCGCGTAGCGGAGACACCGGGCACAGCGCTGTCTCAAGAGAGCTGCACACGCGCGAGCC	421

QY 121 ATGLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 422 CGGCTGGGGCGGACATGAGACGTCGTGGCCCGCTGTGCTACGCGCGGAGG 481
 QY 141 GlnAlaMetLeuAspGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 482 CAGGCGCATCTCGGCGACAGCACCGAGGAGCTGGGTGCGCTGCCCTCCACCTCGCGC 541
 QY 161 LysLeuArgValArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 |||||||
 Db 542 AAGCTGCGGAGCGGCTCTCCGCGATCCCGATGACCTGCAGAACGCGCTGGCAGTGC 601
 QY 181 GlnAlaGlyAlaArgGlyValArgGlyLeuSerAlaIleArgGlyLeuArgGly 200
 |||||||
 Db 602 CAGGCGGGGGCGCGGAGGCGCGCGCGCTCAGCGCCATCCGCGAGCGCTGGG 661
 QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 Db 662 CCCCTGTGTGAGACAGGCGCGCTGGCGCGCCACTGTGTGGCTCCCTGGCGCGCAGCGC 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGluArgLeuArgAlaArgMetGlnLumetGly 240
 |||||||
 Db 722 CTACAGAGAGCGGCGCGAGGCTGGGCGCGCTGGCGCGGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 |||||||
 Db 782 ACTCGGACCGCGGACCGCTGTGACAGGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 |||||||
 Db 842 CTGGAGAGAGCAGCGCCAGCAGATACGCCGCGAGCGCGCTCCAGGCGCGCTCAAG 901
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnIntrPalaGlyLeuValGluLys 300
 |||||||
 Db 902 ACTGTGTGAGCGCGCTGTGAGAGATGACAGCGCCAGTGGCGCGGTGTGTGAGAGAG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||||
 Db 962 GTGACAGGCGCGCGTGGGACCAAGCGCGCGCTGTGTGCCAGGACATATAC 1012

RESULT 12
 HUMAPOE
 LOCUS HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 DEFINITION Human apolipoprotein E mRNA, complete cds.
 ACCESSION M12529
 VERSION M12529.1 GI:178848
 KEYWORDS apolipoprotein.
 SOURCE Homo sapiens (clone: pHAEl112,178,813.) male 57-year old liver
 CDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and
 Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. CDNA cloning and nucleotide sequencing
 of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6337682
 FEATURES
 source location/Qualifiers
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="pHAEl112,178,813."
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 mRNA

CDS
 /gene="APOE"
 /note="apoe mRNA; G00-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:G00-119-691"
 /translation="MKVIMALLVTFLLAGCAKVEQAVETPEPELRQTEKSGQR
 ELAIGREWDYLRWQTSSEYQVEELSSQVEELRALMDETREKELAKVSLERDLP
 VAETRLARLSEKLEOTAGRLGADMEWCGRLVOYRGVQAMIGSTELRYRLSHR
 KLRRLDPDLOKRLAVYOGAREGERSLAIREFLGLVEQGRVRAATVSLAG
 OPLERAOAMEERLRLARMEGSRTRDLDEVKRQVAVRAKLEEQAOQIRLQAEAFQ
 ARLSMPEPLVEDMDQRMAGLVEKVAVGTSAPVPSDNH"
 62..115
 sig_peptide
 /gene="APOE"
 /note="G00-119-691"
 116..1012
 mat_peptide
 /gene="APOE"
 /product="apolipoprotein E"
 /note="G00-119-691"
 370
 variation
 /gene="APOE"
 /note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 410
 variation
 /gene="APOE"
 /note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 569
 variation
 /gene="APOE"
 /note="c in pHAEl112,1781; g in pHAEB13; G00-119-691"
 /replace="g"
 784
 variation
 /gene="APOE"
 /note="t in pHAEl112,1781; c in pHAEB13; G00-119-691"
 /replace="c"
 859
 variation
 /gene="APOE"
 /note="g in pHAEl112,1781; a in pHAEB13; G00-119-691"
 /replace="a"
 BASE COUNT 212 a 370 c 426 g 149 t
 ORIGIN 30 bp upstream of BstNI site.
 Alignment Scores:
 Pred. No.: 2.69e-83 Length: 1157
 Score: 1561.00 Matches: 313
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.24% Indels: 0
 Gaps: 0
 DB: 9
 US-09-827-854-17 (1-317) x HUMAPOE (1-1157)
 QY 1 MetLysValLeuThrPalaAlaLeuValITrrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 62 ATGAAGCTTCTGTGGCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
 |||||||
 Db 122 GAGCAAGCGGTGGAGACAGACGCGGAGCGCGGAGCTGCGCACAGACGACGAGCGAGAC 181
 QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 Db 182 GGCCAGCGCTGGGAGCTGACATGCGGTCTTTTGGGATACCTCGCGGTGGTGCAGACA 241
 QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||||
 Db 242 CTGTCTGACAGAGAGAGGAGCTGCTCAGCTCCCAAGTACCAAGAACTGAGGGG 301
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100

|||||
Db 302 CTATGAGACAGACCAATGAAAGAGTTGAMAGCCCTACAAATCGAAGCTGGGAGAACACTG 361
QY 101 TheProVa1A1aG1uG1uThrArGa1AaRgLeuSerLySG1uG1uG1uA1aG1uA1a 120
|||||
Db 362 ACCCGCGTACGAGAGAGAGAGCGGGGACGCGCTCTCCAGAGAGCTGCAGACGGCGAGGCC 421
QY 121 ArgLeuG1yA1aAspMetG1uAspValCysG1yArgLeuValG1nTyrArgG1yG1uVal 140
|||||
Db 422 CGGCTGGGCGCGGACATGAGAGACGTGTGGCGCGCCCTGTGTCAGTACCCCGCGAGAGTG 481
QY 141 G1uA1aMetLeuAspG1nSerThrG1uG1uLeuArgValArgLeuA1aSerHisLeuArg 160
|||||
Db 482 CAGGCCATGCTCGCGACAGACACCGAGAGCTCGGGTGGCTCGCTCCACCTGCGC 541
QY 161 LysLeuArgLyArGLeuLeuArGAspA1aAspAspLeuG1nLysCysLeuA1aVal1Tyr 180
|||||
Db 542 AAGCTGCGTAAAGCGGCTCTCCCGCATCCCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 601
QY 181 G1uA1aG1yA1aAaRgG1uG1yA1aG1uA1aRgG1yLeuSerA1a1leAaRgG1uAaRgLeuG1y 200
|||||
Db 602 CAGCGCGGGGCGCGGAGAGGCGCGGAGCGGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
QY 201 ProLeuValG1uG1uG1uArgValAaRgA1a1aThrValG1ySerLeuA1aG1uG1nPro 220
|||||
Db 662 CCCCTGTGTGAACAGAGCGCGCGTGGCGCGCCACTGTGGCTCCCTGGCGCGCGCAGCGC 721
QY 221 LeuG1uG1uAaRgA1aG1uA1aTrpG1yG1uAaRgLeuAaRgA1aAaRgMetG1uG1uMetG1y 240
|||||
Db 722 CTACAGAGAGGGGCGCCAGCGCTGGGGGAGCGCGCTGCGCGCGGATGAGAGATGGGG 781
QY 241 SerArG1ThrArGAspArGLeuAspG1uValLysG1uG1uValA1aG1uA1aValA1aLys 260
Db 782 AGTCGAGACCGCGCACCGCTCGGACGAGTGAAGAGCAGAGTGGCGGAGGCGCGCGCAAG 841
QY 261 LeuG1uG1uG1uA1aG1uG1uG1uA1aRgLeuG1uA1aG1uA1aPheG1uA1aAaRgLeuLys 280
Db 842 CTGAGAGAGAGCGCCACAGCATACCGCTGCAGAGCGCGAGCGCTTCCAGCGCGCGCTCAAG 901
QY 281 SerTrpPheG1uProLeuValG1uAspMetG1uAaRgG1nTrpA1aG1yLeuValG1uLys 300
Db 902 AGCTGTGTGAGCGCCCTGGTGGAAAGACATGACAGCGCGCAGTGGCGGGCTGGTGGAGAG 961
QY 301 ValG1uA1aA1aValG1yThrSerA1aA1aProValProSerAspAsnHis 317
Db 962 GTGCAGAGTGCCTGTGGGACACGACGCGCCCTGTGCGCAGCGACATCAGC 1012
RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279 Apo E humanized mammal.
DEFINITION
ACCESSION BD004279.1 GI:18632240
VERSION BD004279.1 GI:18632240
KEYWORDS JP 2001017028-A/3.
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1156)
Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
Apo E humanized mammal
Patent: JP 2001017028-A 3 23-JAN-2001.
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PE 28-Apr-2000 JP 2000128919
PR SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09///C07K14/775,
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC

Alignment Scores:
Pred. No.: 4,03e-83 Length: 1156
Score: 1558.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.05% Indels: 0
Gaps: 0
US-09-827-854-17 (1-317) x BD004279 (1-1156)

QY 1 MetLysValLeuTrpA1a1aLeuLeuVal1ThrPheLeuA1aG1yCysG1nA1aLysVal 20
Db 61 ATGAAGTTCTGTGGGCTCGTTGCTGTGCATTCCTCGCAGAGATGCCAGGCAAGGTG 120
QY 21 G1uG1uA1aValG1uThrG1uProG1uProG1uLeuArgG1nThrG1uTrpG1nSer 40
Db 121 GAGCAAGCGGTGAGACAGACGCGGAGCGCGAGCTCGCGCACAGACAGCGAGTGGCAGAC 180
QY 41 G1yG1uAaRgTrpG1uLeuA1aLeuG1yAaRgPheTrpAspTyrLeuAaRgTrpValG1nThr 60
Db 181 GCGCACGCTGGGAACCTGGCACCTGGTCTTTGGGATTACCTGCGTGGGTGAGACA 240
QY 61 LeuSerG1uG1uValG1nG1uLeuLeuSerSerG1nVal1ThrG1nG1uLeuAaRgA1a 80
Db 241 CTGTCTGACAGAGTGTGAGAGAGAGCTGCTCACCTCCAGTCAACCAGAACTGAGAGCGG 300
QY 81 LeuMetAspG1uThrMetLysG1uLeuLysA1aTyrLysSerG1uLeuG1uG1uG1nLeu 100
Db 301 CTGATGACAGACCACTGAAGAGTTGAAGGCTTACAAATTCGAACTGAGAGAACACTG 360
QY 101 TheProVa1A1aG1uG1uThrArGa1AaRgLeuSerLySG1uG1uA1aG1uA1a 120
Db 361 ACCCGCGTGGCGAGAGACGCGGCGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGCAGGCC 420
QY 121 ArgLeuG1yA1aAspMetLysAspValCysG1yArgLeuValG1nTyrArgG1yG1uVal 140
Db 421 CGGCTGGGCGCGACATGAGAGACGTGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 141 G1uA1aMetLeuAspG1nSerThrG1uG1uLeuArgValArgLeuA1aSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCGACAGACACCGAGAGCTGCGGCTGCGCTGCGCTGCGCTGCGCTGCGC 540
QY 161 LysLeuArgLyArGLeuLeuArGAspA1aAspAspLeuG1nLysCysLeuA1aVal1Tyr 180
Db 541 AAGCTGCTAAAGGGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCTGCAGAGTAC 600
QY 181 G1uA1aG1yA1aAaRgG1uG1yA1aG1uA1aRgG1yLeuSerA1a1leAaRgG1uAaRgLeuG1y 200
Db 601 CAGGCCGCGGCGCGGAGAGCGCGGAGCGCGCTCAGCGCATCCGCGAGCGCTGGGGG 660
QY 201 ProLeuValG1uG1uG1uArgValAaRgA1a1aThrValG1ySerLeuA1aG1uG1nPro 220
Db 661 CCCCTGTGTGAACAGAGCGCGCTGGCGCGCGCAGCTGTGGCTCTGCGCGCGCGCGCGCG 720
QY 221 LeuG1uG1uAaRgA1aG1uA1aTrpG1yG1uAaRgLeuAaRgMetG1uG1uMetG1y 240
Db 721 CTACAGAGAGCGGCGCGAGCGCTGGGCGAGCGCGCTGCGCGCGGTGAGAGATGGGC 780
QY 241 SerArG1ThrArGAspArGLeuAspG1uValLysG1uG1uValA1aG1uA1aValA1aLys 260
Db 781 AGCGGAGCCCGGACCGCTGTGAAGAGAGTGAAGAGCAAGTGGCGAGTGGCGCGCAAG 840
QY 261 LeuG1uG1uG1aG1uG1uG1uA1aRgLeuG1uA1aPheG1uA1aAaRgLeuLys 280
|||||

Db 841 CTGGAGGACGAGCCAGCATACGCTTCAGAGCCGCTCCAGGCCCTCCAG 900
QY 281 SerTpphcg1uProleuValGluaspMetGlnArgGlnTrpAlaGlyLeuValGluys 300
Db 901 ACTGTTGACCCCTGGTGGAGACATGACGCGCACTGGGCGGCTGGGAGAG 960
QY 301 ValGlnAlaAlaValAlGlyThrsSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTCAGAGCTGCGGTGGGACGAGCGCGCCCTGTGTCCAGAGACATAC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS Monkey mRNA for apolipoprotein E.
DEFINITION X13887
ACCESSION X13887.1 GI:38054
VERSION apolipoprotein; apolipoprotein E.
KEYWORDS Macaca fascicularis.
SOURCE Macaca fascicularis.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti K.R., Whitted, B.E., Castle, C.K., Polites, H.G. and
Meichlor, G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source Location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_lib="lambda gt11"
83..1036
/codon_start=1
/product="apolipoprotein E (AA -18 to 299)"
/protein_18="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SMISS-PROT:P10517"
/translation="MKVLMALIVTFLGCOAKYQEPVETPELROAGSGSPM
ELALGRFMDYLRWQOTLSEVOEELSPVOETLTMDTKELKAKSKFLPDLSP
VAETRRARLSLEIOAARIGADMEDVSRIVQYRSFVQAMLGSTEBELRRLASHLR
KLKRLRLDADLDKRLAVYQAGABEGAEVSAIREFLGLVEGGRVRAIVGSLAS
QPLDERAQLGERLKAARMEGSRTRDLDEKQVAEVRALKEQAOQISLQAEAFQ
ARLSKSFPEPLVEDMKRWAGLVEKVOAAGVASTAPVPIDNH"
83..136
sig_peptide 137..1033
mat_peptide /product="apolipoprotein E (AA 1 - 299)"
1154..1159
misc_feature /note="polyA signal"
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 5.67e-76 Length: 1178
Score: 1436.00 Matches: 294
Percent Similarity: 93.38% Conserved: 2
Best Local Similarity: 92.74% Mismatches: 21
Query Match: 90.37% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-17 (1-317) x MFAPOE (1-1178)
QY 1 MetlyValLeuTrpAlaAlaLeuLeuValThrpheLeuAlaGlyCysGlnAlaLysVal 20

Db 83 ATGAAGTTCGTGGGCTCGTTGCTGTCACATTCCTGGCAGATGCCAGGCAAGTG 142
QY 21 GluGlnAlaValAlGluThruLeuProGluProGluLeuArgGlnGlnThruTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 203 GGCACGCTGGGAGCTGGACACTGGGTCCCTTTTGGATTAACCTGCGTGGTGACAGA 262
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 263 CTGTCTGACAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 322
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
Db 323 CTGATGGAGACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 382
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 383 ACCCGGTGGGAG 442
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 443 CGGTGGGTGGTCCGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 502
QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCCATCTGTGGGCGAGTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db 563 AAGCTGCGCAACCGGCTCTCCGCGATGCTGATGACCTCACAAGAGCGCTGGCAGTGT 622
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 622 CAGCGCGGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
Db 683 CCCCTGTGAGACAG 742
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 743 CTTCAGGAGAGCGGCGCAGGCGCTTGGGTGAGCGGCTTGGCGCACGGATGAGAGAGTGGGC 802
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 803 ACCCGAGAGCGGAG 862
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 863 CTGGAGGAACAGAGCCAGAGATAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 922
QY 281 SerTpphcg1uProleuValGluaspMetGlnArgGlnTrpAlaGlyLeuValGluys 300
Db 923 ACTGTTGACCCCTGGTGGAGATATGACGCGCCAGTGGGCTGGTGGAGAG 982
QY 301 ValGlnAlaAlaValAlGlyThrsSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTCAGAGCTGCGGTGGGACGAGCGCGCCCTGTGTCCAGAGACATAC 1033
RESULT 15
AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS Homo sapiens apolipoprotein-E gene, complete cds.
DEFINITION AF261279
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salomaa,V., Boerwinkle,E. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
apolipoprotein E gene
JOURNAL Genome Res. 10 (10), 1532-1545 (2000)
MEDLINE 20499366
PubMed 11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
Location/Qualifiers
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
repeat_region <3..>108
/note="putative"
/rpt_family="MTR"
/rpt_type="dispersed"
variation 73
/frequency="0.01"
/replace="t"
repeat_region <207..>295
/note="putative"
/rpt_family="MTR"
/rpt_type="dispersed"
variation 308
/frequency="0.01"
/replace="t"
satellite <320..>339
/note="putative"
/rpt_type="tandem"
repeat_region <340..>637
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 471
/frequency="0.01"
/replace="g"
variation 545
/frequency="0.01"
/replace="t"
variation 560
/frequency="0.22"
/replace="t"
variation 624
/frequency="0.07"
/replace="c"
satellite <638..>718
/note="putative"
/rpt_type="tandem"
variation 832
/frequency="0.45"
/replace="t"
mrna join(1060..1094,1855..1920,3013..3205,3786..4645)
/product="apolipoprotein-E"
variation 1163
/frequency="0.35"
/replace="c"
variation 1522
/frequency="0.01"
/replace="a"
variation 1575
/frequency="0.01"
/replace="t"
join(1878..1920,3013..3205,3786..4503)
/note="APOE"
/codon_start=1

/product="apolipoprotein-E"
/protein_id="AAG27089.1"
/db_xref="GI:11034801"
/translation="MKVLMALILYFLACGCAKVEQAVETPEPELROOTEMSGORM
ELALGRFMDYLRWVQTLSEQVDELLSSGYTOBLRLANDETKELKATKSELEBQITP
VAETVARLSKELQAQAARLQADMEVCGLVQYRGEVAMLGQSTBELRLASHLR
KLRLRLRDADLQRLAVYQAGARGAGRGISAJRERGLPVEQGRVRAAVGSLAG
QPLQERAQAGGERLARMEKMSRTDRIDEVKEQVAVAEVRATLEEQAOOIRLQAEAFQ
ARLKSWEPEPLVEDMQRONAGLVEKVOAAVGTSAAPVPSDNH"
variation 1998
/frequency="0.10"
/replace="a"
repeat_region <2124..>2435
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 2440
/frequency="0.21"
/replace="a"
repeat_region <2569..>2848
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 2907
/frequency="0.01"
/replace="g"
variation 3106
/frequency="0.01"
/replace="c"
repeat_region <3472..>3588
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 3673
/frequency="0.01"
/replace="g"
variation 3701
/frequency="0.01"
/replace="c"
variation 3937
/frequency="0.14"
/replace="c"
variation 4036
/frequency="0.01"
/replace="t"
variation 4075
/frequency="0.07"
/replace="t"
repeat_region <4755..>5056
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation <5065..>5476
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
variation 5229
/frequency="0.03"
/replace="g"
variation 5229
/frequency="0.07"
/replace="t"
variation 5229
/frequency="0.40"
/replace="g"
variation 5230
/frequency="0.13"
/replace="a"
variation 5361
/frequency="0.06"
/replace="c"
BASE COUNT 1029 a 1663 c 1603 g 1196 t
ORIGIN

Alignment Scores:

Pred. No.:	4,37e-72	Length:	5491
Score:	1381.50	Matches:	302
Percent Similarity:	60.408	Conservative:	0
Best Local Similarity:	60.408	Mismatches:	4
Query Match:	86.94%	Indels:	194
DB:	9	Gaps:	1

US-09-827-854-17 (1-317) x AF261279 (1-5491)

```
QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
    |||
Db 3003 TTCACACAGAGATGCCAGGCCCAAGGTGGAGCAAGCGGTGGAGACAGAGCCGAGCCCGAG 3062
QY 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuLysArgPhe 51
    |||
Db 3063 TGCCGCGACAGACCGAGGCGAGCGCGCCAGCGCTGGGAACCTGGCACTGGGCTGCTTT 3122
QY 52 TrpAspTyrLeuArgTrpValGlnThrLeuSerGluGlnValGlnGluLeuLeuSer 71
    |||
Db 3123 TGGGATTAAGTCTGCTGGGTGACAGACACTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGC 3182
QY 72 SerGlnValThrGlnGluLeu----- 78
    |||
Db 3183 TCCACAGTCACCCAGAGAACTGAGTGAAGTCCCATCCGCGCTGGCCCTTGACCCCTCGTG 3242
QY 78 ----- 78
Db 3243 GCGGCGTATACCTCCCAAGTCCAGGTCAGATTTCATTCGCCCTGTGCTTGGGGGG 3302
QY 78 ----- 78
Db 3303 CCTGGGTCTCTGCTGTTCTTACTTCTCTTCCCATTTCTGACTCCTGGCTTTAGCTTC 3362
QY 78 ----- 78
Db 3363 TGGAAATTCCTCTCTCAGCTTGTCTCTCTCTTCCCTTCCAGTCAAGTCTCTCAGACT 3422
QY 78 ----- 78
Db 3423 CGTCTGGCTGTCTGTCTGCTCCCTTACCTTTTATATAGAGACAGAGATGGGGT 3482
QY 78 ----- 78
Db 3483 CTCACGTGTGTGCCAGGCTGTGTAACCTCTGGGCTCAAGCGATCCTCCGCCCTCG 3542
QY 78 ----- 78
Db 3543 CCTCCCAAGTGTGGGATTAGAGCATAGCCACCTTGGCCGCGCTCTACTCTCTTCT 3602
QY 78 ----- 78
Db 3603 TCGTCTGACCTGACCTCTGATCTGTCTGATCTGTCTGTCTGCTCTCTCTCTG 3662
QY 78 ----- 78
Db 3663 GCGCTGCGCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3722
QY 78 ----- 78
Db 3723 CCCCATCCAGCCCTTTCCTCCCGCTCCCACTGTGCGACACCCCTCCGCGCTCTGCGCG 3782
QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnG 98
    |||
Db 3783 CAGGCGCGTGAAGAGACCATGAAAGAGTTGAAGGCTTACCAATCGGAACCTGGAGGA 3842
QY 98 uGlnLeuThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAl 118
    |||
Db 3843 ACAACTGACCCGCGTGGCGGAGAGACCGCGGCTGTCCAAAGAGAGCTCAGGCGCG 3902
QY 118 aGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgG 138
    |||
```

```
Db 3903 GCAGCCCGGCTGGCGCGGACATGAGAGACGTGTGCGGCCCTGTGTGCTACTACCGCG 3962
QY 138 YGluValGlnAlaMetLeuAspGlnSerThrGluGlnLeuArgValArgLeuAlaSerHi 158
    |||
Db 3963 CGAGGTGACAGGCTATGCTTGCGCCAGAGACCCAGAGAGCTGCGGGTGTGCTGCTGCCA 4022
QY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAl 178
    |||
Db 4023 CCTGCGCAAGCTGCGTAAAGCGGCTCTCCGCGATGCCATGTACCTGCAGAAACGCGCTGC 4082
QY 178 aValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluAr 198
    |||
Db 4083 AGTTTACACAGGCGGGGGCGCGAGAGCGCCAGAGCGGCTTCACAGCGCATCCGCGAGCG 4142
QY 198 gLeuGlyProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaG 218
    |||
Db 4143 CTTGGGGCCCTGTGTGAGACAGGCGCGCTGCGGGCGGCTGCTGTGGCTCTGCTGCGCG 4202
QY 218 YGlnProLeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGlu 238
    |||
Db 4203 CCAGCGGCTACAGAGCGGGGCCAGAGCTGCGGGCGAGCGGCTGCGCGCGGATGGAGGA 4262
QY 238 uMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAr 258
    |||
Db 4263 GATGGGACAGCCGACCCCGACCGCTGGACGAGGTGAAGAGCAGGTGGCGAGGTGCG 4322
QY 258 gAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaAr 278
    |||
Db 4323 CGCCAAAGCTGGAAGAGAGGCGCCAGCATGAGCTGCGAGGCGGAGGCTTCCAGGCGCG 4382
QY 278 gLeuLysSerTrpPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuVa 298
    |||
Db 4383 CCTCAAGAGCTGTGTGAGACCCCTGTGTGAGACATGACAGCCCATGTGGCGGCGGTGT 4442
QY 298 IGlLysValGlnAlaAlaValGlyThrSerAlaIleProValProSerAspAsnHis 317
    |||
Db 4443 GGAGAAGGTGAGGCTGCGGTGGCGACAGCCGCGCTGTGGCCAGGACATAC 4500
```

Search completed: March 14, 2003, 17:32:36
Job time : 2314.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 178.728 Seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-17

Perfect score: 1589

Sequence: 1 MKVLMAALVFLAGCAK.....VEKVAAGTSAAPVSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/SPFO/spool/US09827854/turnat_11032003_101609_27476/app.query.fasta.1.3576
-DB=N.Geneseq_101002 -OFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum2 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854 -CGCN.1.1.1201 -turnat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO.XIPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LOGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-WGAPOP=10 -WGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_101002: *
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	24	AAD22050	Human apolipoprotein
2	1582	99.6	1156	22	AAE84314	Human ApoE2 coding
3	1582	99.6	1156	24	AAD22049	Human apolipoprotein
4	1570	98.8	954	24	AAD26035	Human apolipoprotein
5	1570	98.8	1110	7	AA606409	Human apolipoprotein
6	1570	98.8	1147	22	ABA83113	Apolipoprotein E o
7	1570	98.8	1156	22	AAF84315	Human ApoE3 coding
8	1570	98.8	1156	24	AAD22048	Human apolipoprotein
9	1566	98.6	1156	24	AAD22052	Human apolipoprotein
10	1562	98.3	1156	24	AAD22051	Human apolipoprotein
11	1561	98.2	1157	17	AAT06957	Human apolipoprotein
12	1561	98.2	1157	24	ABN95746	Gene #2244 used to
13	1561	98.2	1157	24	ABK64514	Human benign prost
14	1561	98.2	1157	24	ABL65450	Lung cancer relate
15	1560	98.2	1110	6	AAN50450	Sequence encoding
16	1558	98.0	1156	22	AAF84316	Human ApoE4 coding
17	1558	98.0	1156	24	AAD22047	Human apolipoprotein
18	1466.5	92.3	1279	22	AA522437	Human cDNA encodin
19	1447.5	91.1	1107	19	AA75756	Human apolipoprotein
20	1381.5	86.9	9360	24	ABL31915	Human ApoE genomic
21	1381.5	86.9	10716	24	AAD26034	Human apolipoprotein
22	1376	86.6	3805	20	AA209524	Human Apo E genomic
23	1376	86.6	3805	20	AA209526	Human Apo E genomic
24	1361.5	85.7	10716	24	AAD26108	Human Apo E genomic
25	1148.5	72.3	965	24	AAD32081	Human alpha-1-anti
26	1148.5	72.3	5617	24	AAD32077	Human albumin prom
27	1148.5	72.3	6026	24	AAD32075	Nucleotide sequenc
28	1123	70.7	1126	19	AAV28159	ApoE4Lx2 protease
29	970	61.0	936	15	AAQ69101	Human ApoE4Lx2 CDN
30	970	61.0	936	17	AAT18070	Partial human apol
31	969	61.0	660	18	AAT69792	Human ApoE4 CDNA.
32	932	58.1	597	17	AAT18068	ApoE4L protease cd
33	923	58.1	597	15	AA609099	Human cDNA encodin
34	869	54.7	1381	22	AA522673	Human cDNA associ
35	852.5	53.7	786	21	AAF18114	EST clone CP147.
36	663	41.7	600	20	AAV89595	Human cDNA used to
37	651	41.0	407	24	ABK34238	Colon adenocarcino
38	625	39.3	478	24	ABL62679	Thyroid cancer rel
39	625	39.3	478	24	ABL67340	Human diagnostic a
40	625	39.3	478	24	AA503049	Human secreted pro
41	594	37.4	499	22	AA502139	Human apolipoprote
42	504	31.7	405	21	AAQ11980	Human Est-derived
43	499	31.4	330	12	AAQ11980	Human Est-derived
44	485	30.5	345	22	AAH98479	Mouse secreted exp
45	410	25.8	382	21	AAA44191	

ALIGNMENTS

RESULT 1
AAD22050 standard: DNA; 1156 BP.

AAD22050: 12-FEB-2002 (first entry)

Human apolipoprotein E (apoE) allele, apoE1 DNA.

Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;

hypertriglyceridaemia; low density lipoprotein; LDL; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 61..1014
/*tag= a

FT sig-peptide /product= "Human apoE allele, apoE1"
FT 61..114 /*tag= b
FT mat-peptide 115..1011 /*tag= c
FT /product= "Mature human apoE allele, apoE1"
XX
PN W020017136-A1.
XX
XX 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
XX
PR 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
XX WPI: 2002-010885/01.
DR P-PSDB; AAE13296.
XX
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridemia -
XX
XX
PS Claim 14; Page 82; 91pp; English.
XX
XX The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE1 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 209 A; 367 C; 431 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 2,23e-115 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-17 (1-317) x AAD22050 (1-1156)

QY 1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlycSerGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGCTGGTGTGGTCACATTCCTGGCAGAGGCCAGCGGAG 120
QY 21 GUGGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GAGCAAGCGGTGAGACAGACGCCGAGCTCGCCAGCAGACCCGAGTGGCAGAGC 180
QY 41 G1yGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCCAGCGCTGGGAACTGGCAGCTGGCTTTGGGATTACTGGCTGGCTGGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAlaGlnAla 80
DB 241 CTGTCTGACGACGAGTGAAGAGAGGCTGCTCCAGCTCCAGGATCGAGAGGCG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGAGACGACCATGAAGAGTTGAAGCGCTCAAAATCGGAATCGAGAGCAACTG 360

QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 361 ACCCGGTGGCGAGGAGAGCGCGGCACGCTCTCCAGAGCTGCAGCGCGCAGGCC 420
QY 121 ArgLeuGlnValAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyVal 140
DB 421 CGGCTGGCGCGGACATGAGGACGTGTGGCGCGCTGGTGCAGTACCGCGGAGGAG 480
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGCACACAGACACGAGAGCTCGGGGTGGCTCCCTCCCTCCAGTGGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
DB 541 AAGCTCGTAAAGGCGCTCCCGCATGCCGATGCCGATGACCTGCGCAGTGTAC 600
QY 181 GlnAlaGlnValArgGlnGlnValArgGlyLysSerAlaIleArgGlnArgLeuGly 200
DB 601 CAGCGCGGGCGCGCAGAGGCGCGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGTGGAAACAGGCGCGCGTGGCGCGCCACTGTGGCTCCCTGGCGCGCCAGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrGlyLysArgLeuArgAlaArgMetGlnMetGly 240
DB 721 CTACAGACAGCGGCGCCAGCGCTGGCGAGCGGCTCGCGCGGATGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 781 AGCGGAGCGCGGACCGCGCTGGCAGAGGTGAAGACAGCTGGCGAGCTGGCGCGCAAG 840
QY 261 LeuGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGAGAGACAGGCGCCAGCAGATACCTCTGCAGCGCGAGCGCTCCAGCGCGCTCAAG 900
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
DB 901 AGCTGTTCGAGGCGCTGTGTGAAGACATGACAGCGCCAGTGGCGCGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGTGGTGGCACCAGCGCGCCCTGTGCTCCAGCAGCAATAC 1011

RESULT 2
AAF84314
ID AAF84314 standard; cDNA; 1156 BP.
XX
AC AAF84314;
XX
XX 21-JUN-2001 (first entry)
DT
XX
XX Human ApoE2 coding sequence.
DE
XX Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 61..1014
FT /tag= a
FT /product= "Human ApoE2"
PN JP2001017028-A.
XX
XX 23-JAN-2001.
PD
XX
XX 28-APR-2000; 2000JP-0128919.
PF
XX
XX 06-MAY-1999; 99JP-0125647.
PR
XX
XX (MITU) MITSUBISHI CHEM CORP.
PA


```
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||||
Db 421 CCGCTGGCCCGGACATGAGGAGACGTGTGGCCGCTGGTGCAGTACCCCGGAGAGTGT 480
QY 141 GluAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 481 CAGGCCATGTCTCGGCGACGACCGAGGAGCTGGGGTGGCTCCGCTCCCACTTCGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
    |||||||
Db 541 AAGCTGGCTAAGCGGCTCCCTCCGGATGCCGATGACTGCAGAAAGCCCTGGCAGTGTAC 600
QY 181 GluAlaGlyAlaArgGluGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||||
Db 601 CAGCGCCGCGGCGCCGAGGCGCGGAGCGGCGCTCAGCCCATCCCGAGCGCCTGGGG 660
QY 201 ProLeuValGluGluGlnLysArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
    |||||||
Db 661 CCCCTGGTGGACAGGCGCCCGCTGGGCGGCCGCTGTGGCTCCCTGGCGGCGCAGCG 720
QY 221 LeuGluGlnLysArgAlaGlnAlaTyrPglyLysArgLeuArgAlaArgMetGluGluMetGly 240
    |||||||
Db 721 CTACAGGAGAGCGGCGCCAGCGCTGGGCGAGCGGCTGGCGCGGATGGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||||
Db 781 AGCGGACCGCGCAGCCCTGGAGAGGTAGAGAGCAGGTGGCGGAGTGGCGGCCAAG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaIleArgLeuLys 280
    |||||||
Db 841 CTGGAGAGAGAGCGCCCGCAGATACGCTCTGAGGCGCAGGCGCTTCCAGGCGCCGCTCAAG 900
QY 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuValGluLys 300
    |||||||
Db 901 AGCTGTGTTCCAGGCCCTCGGTGGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
QY 301 ValGluAlaIleValAlaGlyThrSerAlaIleProValProSerAspAsnHis 317
    |||||||
Db 961 GTGCAGCTGCTCGTGGGACACGCGCCGCTGTGCCAGCAATCATC 1011
RESULT 8
AAD22048
ID AAD22048 standard; DNA: 1156 BP.
AC AAD22048:
XX
XX
DT 12-FEB-2002 (first entry)
XX
DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
XX
XX Human apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT FT /product= "Human apoE isoprotein, apoE3"
FT FT /tag= a
FT FT /tag= b
FT FT /tag= c
FT FT /product= "Mature human apoE isoprotein, apoE3"
XX
XX
PN WO200177136-A1.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-US11358.
XX
XX 06-APR-2000; 2000US-0544386.
XX
PR
```

```
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
XX Zannis VI, Kypros KE;
PI
XX
DR WPI: 2002-010885/01.
DR P-PSDB: AAE13294.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia.
XX
PS Claim 14: Page 81; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
XX
Alignment Scores:
Pred. No.: 6, 81e-114 Length: 1156
Score: 1570.00 Matches: 315
Percent Similarity: 99.378 Conservative: 0
Best Local Similarity: 99.378 Mismatches: 2
Query Match: 98.80% Indels: 0
DB: Gaps: 0
US-09-827-854-17 (1-317) x AAD22048 (1-1156)
QY 1 MetLysValLeuTyrPalaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
Db 61 ATGAGAGTTCTGTGGGCTGCTGTGTGTCATCTTCGTCAGAGATGCCAGGCGCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrPalaGlnSer 40
    |||||||
Db 121 GAGCAAGCGGTGGAGACAGAGCGCGGAGCCGAGCTCGCCAGACAGCCAGTGGCAGAC 180
QY 41 GlyLysArgTyrPalaGluLeuAlaLeuGlyArgPheTyrAspTyrLeuArgTyrPalaGlnThr 60
    |||||||
Db 181 GGCACGCGTGGGAGACTGGAGTGGGTCCCTTTGGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGluGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
Db 241 CTGTGACAGAGGTGACAGAGAGAGCTGTGCTCAGCTCCAGATCACCCAGAACTGAGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
    |||||||
Db 301 CTGATGACAGAGACCATGGAAGGAGTTGAAGGCCCTTCAAAATCGGAACCTGAGGAACATCG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
    |||||||
Db 361 ACCCGGTGGCGGAGAGAGAGCGGCGCAGGCTGTCCAAAGAGCTGAGAGCGGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||||
Db 421 CCGCTGGCGCGGACATGAGGAGCGTGTGGCCGCTGGTGCAGTACCCCGGAGAGTGT 480
QY 141 GluAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 481 CAGGCCATGTCTCGGCGACGACCGAGGAGCTGGGGTGGCTCCGCTCCCACTTCGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
    |||||||
```

Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCGATGACTGACAGAGCGGCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnuArgGlyLeuSerAla11LeArgGlnuArgLeuGly 200
Db 601 CAGGCGGGGGCCCGGAGGGCCGAGCGGCGCTCAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGCTGTAACAGAGCGGCTGGCGGCGGCGGCTGTGGCTCTGGCCGCGCAAGCCG 720
QY 221 LeuGlnGlnuArgAlaGlnAlaThrPheGlyuArgLeuArgAlaArgMetGlnuGly 240
Db 721 CTACAGAGCGGCGGCGGCGGCTGGCGGCGGCTGGCGGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnuValGlyGlnGlnuValAlaGlnuValArgAlaGly 260
Db 781 ACCCGGACCGCGGACCGCTGGAGCGAGTGAAGAGAGAGAGTGGCGAGGTGGCGCCCAAG 840
QY 261 LeuGlnGlnuArgAlaGlnGlnuArgLeuGlnAlaGlnuArgLeuGlnuValArgLeuGly 280
Db 841 CTGAGAGAGAGAGAGAGAGAGATGAGCGCTGGAGCGGCGGCTTCCAGGCGGCTCAAG 900
QY 281 SerTrpPheGlnuArgLeuValGlnuAspMetGlnuArgGlnuArgAlaGlyLeuValGly 300
Db 901 ACCTGCTTCGAGCGGCTGGTGAAGACATGACAGCCGAGTGGCGGCTGGTGAAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGACAGGCTGGCGGAGCAGCGCGGCTGTGCTCCAGGACATCAC 1011
RESULT 9
AAD22052 standard; DNA; 1156 BP.
ID AAD22052 standard; DNA; 1156 BP.
AC AAD22052;
XX 12-FEB-2002 (first entry)
XX Human apolipoprotein E (apoE) allele, apoE2** DNA.
DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 61..1014
FT /tag= a
FT /product= "Human apoE allele, apoE2**"
FT sig_peptide 61..114
FT /tag= b
FT mat_peptide 115..1011
FT /tag= c
FT /product= "Mature human apoE allele, apoE2**"
PN WO200177136-A1.
XX 18-OCT-2001.
PD
XX 06-APR-2001; 2001WO-US11358.
PF
XX 06-APR-2000; 2000US-0544386.
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX WPI: 2002-010885/01.
DR P-PSDB; AAE13298.
XX

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia
PS Claim 14; Page 83; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2** DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other:
XX
Alignment Scores:
Pred. No.: 1..4e-113 Length: 1156
Score: 1566.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 98.55% Indels: 0
DB: Gaps: 24
US-09-827-854-17 (1-317) x AAD22052 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGGCTGCGTGGTGCATTCCTGGAGAGATGCCAGGCCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnuArgGlnuArgGlnuArgGlnuArgGlnuArgGlnuArg 40
Db 121 GAGCAAGCGGTGAGACAGAGCGGCGGCGGCTGGCGGCGGATGAGAGAGAGAGAG 180
QY 41 GlyGlnArgTrpGlnuLeuAlaLeuGlnuArgPheTrpAspTrpLeuArgTrpValGlnuTr 60
Db 181 GGCCAGCGCTGGGAACCTGGCACTGGGCTTTTGGGATTAACCTGCGTGGGTGCAGACA 240
QY 61 LeuSerGlnuGlnuValGlnuGlnuLeuLeuSerSerGlnuValThrGlnuLeuArgAla 80
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGAGTCAACCAAGAACTGAGGGG 300
QY 81 LeuMetAspGlnuThrMetLysGlnuLeuLysAlaTrpLysSerGlnuGlnuGlnuGlnu 100
Db 301 CTGATGAGCAGACCATGTAAGAGTTGAAGGCTTCAAAATCGAACTGGAGAGAACTACTG 360
QY 101 ThrProValAlaGlnuGlnuThrArgAlaArgLeuSerLysGlnuGlnuAlaAlaGlnuAla 120
Db 361 ACCCGGTGGCGGAGAGAGAGCGGCGGCGGCTGTCCAGAGACTGACGGCGGCGAGGCC 420
QY 121 ArgLeuGlnuAlaAspMetGlnuAspValCysGlnuArgLeuValGlnuTrpArgGlyGlnuVal 140
Db 421 CGGCTGGCGGCGGAGCATGAGAGAGCTGTGCGGCGGCTGTGCTCACTACCGCGGCGAGGTG 480
QY 141 GlnAlaMetLeuAspGlnuSerThrGlnuGlnuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCCAGAGCACCGAGGAGCTGGGGGTGGCTGGCTCCCACTCGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspPheGlnuGlnuGlnuGlnuAlaValTrp 180
Db 541 AAGCTGCTGACAGGCTCTCCGCGATGCGATGACTGACAGAGAGCTGGCACTGTAC 600
QY 181 GlnAlaGlnuArgGlnuGlnuArgGlnuArgGlnuArgGlnuArgGlnuArgGlnuArg 200
Db 601 CAGGCGGGGGCCCGGAGGGCCGAGGCGGCTCAGCGCATCCGAGAGCGGCTGGGG 660
QY 201 ProLeuValGlnuGlnuArgValArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnuPro 220
Db 661 CCCCTGCTGTAACAGAGCGGCTGGCGGCGGCTGTGGCTCTGGCCGCGCAAGCCG 720
QY 221 LeuGlnGlnuArgAlaGlnAlaThrPheGlyuArgLeuArgAlaArgMetGlnuGly 240

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIlys 300
|||||
DB 901 AGCTGGTTCGAGCCCTGGTGGAGACATGCAGCCGACATGGCGCGCTGGTGGAGAG 960
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
DB 961 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCTCCAGCAGCAATCAC 1011
RESULT 11
AAT06957
ID AAT06957 standard; cDNA to mRNA; 1157 BP.
XX
AC AAT06957;
XX
DT 19-JUN-1996 (first entry)
XX
DE Human apolipoprotein-E (ApoE) cDNA.
XX
KW Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;
KW Manduca sexta; Autographica californica nuclear polyhedrosis virus;
KW haemolymph; lipid complex; biologically active; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc_feature 62..119
FT note="misc_signal"
FT CDS 62..1015
FT /tag= b
XX
PN US5472858-A.
XX
PD 05-DEC-1995.
XX
PF 04-JUN-1991; 91US-0709949.
XX
PR 04-JUN-1991; 91US-0709949.
XX
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Attie AD, Beckage NE, Gretsch DG, Sturley SL;
XX
DR WPI; 1996-029812/03.
XX
DR P-PSDB; AAR86791.
XX
PT Prod. of recombinant apo:lipoprotein E in insects - by infecting
PT Manduca sexta larvae with recombinant Autographica californica
PT nuclear polyhedrosis baculovirus vector.
XX
PS Disclosure; Columns 11-14; 10pp; English.
XX
CC Recombinant human apolipoprotein-E (ApoE) (AAR86791) can be produced
CC by preparing a genetic construct (contg. an ApoE-encoding sequence,
CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
CC to be expressed in insect cells), which is then introduced into a
CC Manduca sexta larva (using a recombinant Autographica californica
CC nuclear polyhedrosis virus) and recovering the protein from the
CC haemolymph of the larval host. The ApoE produced is in a form
CC sufficiently complexed with lipids to be biologically active, which
CC cannot be achieved in insect cell cultures, and can therefore be
CC used in therapeutic applications.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 3,44e-113 Length: 1157
Score: 1561.00 Matches: 313
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.24% Indels: 0
DB: 17 Gaps: 0

US-09-827-854-17 (1-317) x AAT06957 (1-1157)
QY 1 MetIysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyIysGlnAlaIysVal 20
|||||
DB 62 ATGAAGTTCTGTGGGCTCGTTGCTGTGCATCTTCCTGGCAGAGATGCCAGGCCAAGG 121
QY 21 GlnGlnAlaValAlaGlyThrGluProGluProGluLeuArgGlnGlnThrGlyIys 40
|||||
DB 122 GAGCAAGCGGTGGAGACACAGCCGAGCCGAGCTCGCCAGCAGACCCAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
DB 182 GCCACAGCTGGGAACTGCGACATGCGCTGCTTGGATTAACCTCGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 242 CTGTCTGACAGAGTGCAGAGGAGCTGTACGCTCCCAAGTCAACCAAGAACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetIysGluLeuIysAlaTrpIysSerGluLeuGlnGlnLeu 100
|||||
DB 302 CTGATGGAGAGACCATGAAGAGGTGAGAGGCTTACAATCGGAACCTGAGAGAACATG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluGlnAlaAlaGlnAla 120
|||||
DB 362 ACCCCGCTAGCGAGAGAGACGCGGCGCTGTCCAAAGAGCTGCAGACGCGCAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValIysGlyArgLeuValGlnThrArgGlyIysVal 140
|||||
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGCGGCGCTGTGTGATGATGACGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 482 CAGGCCATCTCGCGCCAGAGACACCGAGAGCTCGGGCTCGCTGCCCTCCACCTGGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
DB 542 AAGCTGCTGAAGCGGCTCTCCCGCATGCCATGACTCCAGAAAGCGCTGGCAGGTATC 601
QY 181 GlnAlaGlyAlaArgGlnGlnIysAlaGluArgGlyLeuSerAlaAlaArgGluThrArgLeuGly 200
|||||
DB 602 CAGGCGGGGCGCGGAGGCGCGGCGGCTGAGGCCATCCGCGCAGCGCCCTGGGG 661
QY 201 ProLeuValGlnGlnIysArgValArgAlaAlaThrValGlySerLeuAlaGlyIysPro 220
|||||
DB 662 CCCCTGTGTGAACAGGCGCGGCTGCGGCGCCACACTGTGGCTCCCTGGCCGCGCAGCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyIysArgLeuArgAlaArgMetGlnGluMetGly 240
|||||
DB 722 CTACAGAGACCGGGCCAGGCTGTGGGCGAGCGGCTCGCGCGCATGTGAGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValIysGluGlnValAlaGluValArgAlaIys 260
|||||
DB 782 ACTCGGACCCGCGACCGCTTGACGAGGTGAAGAGACAGAGTGGAGGTGGCGCCAGAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuIys 280
|||||
DB 842 CTGGAGAGACAGCCAGAGATACCTCTGAGCGCCGAGAGCTTCAGAGGCCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
|||||
DB 902 AGCTGGTTCGAGCCCTGGTGGAGACATGCAGCCGACATGGCGCGCTGGTGGAGAG 961
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
DB 962 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCTCCAGCAGCAATCAC 1012
RESULT 12
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)

PA (NISB) JAPAN TOBACCO INC.
XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
XX
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
XX cells
XX
PS Disclosure: Page 239-240; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 3 44e-113 Length: 1157
Score: 1561.00 Matches: 313
Percent Similarity: 98.74% Conservatave: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.24% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-17 (1-317) x ABK64514 (1-1157)
OY 1 MetLysValLeuTrpAlaIleuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGGGCTCGCTGCTGTCACATTCCTCGCAGAGATGCCAGGCAAGT 121
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnIns 40
DB 122 GACAGACCGGTGGAGACAGACCGGAGCCGAGACTGCGCAGACGAGCAAGTGCAGAGC 181
OY 41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACAGCGCTGGGAATGGAGCTGCTCTTTGGATTACTGTGGCTGGGAGACA 241
OY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGAGTGCAGAGAGCTGCTCAGTCCCAAGTCAACCAAGACTGAGAGCGG 301
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 302 CTGATGAGAGACCATGAGAGGATTGAAGGCCCTACAAATCGGAACCTGAGAGCAACTG 361
OY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCCGCTGAGGAGAGAGCGGGGCGGCTGTCCAAAGAGCTGAGACGGGCGGCGGC 421
OY 121 ArgLeuGlyAlaAspMetGluAspValGlyArgLeuValGlnTyrArgGlyGlnVal 140

DB 422 CGGCTGGGCGCGACATGAGAGACGTGTGGCGCGCTGTGTGCATGCCGCGGAGGTG 481
OY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGCGCATGCTCGCGCCAGAGACCGAGAGCTCGGGGTGGCTCCCTCCACCTGCGC 541
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 542 AACCTCGTAAAGCGGCTCTCCCGCATCCGATGACTGTCAGAAAGCGCTGCGAGTGTAC 601
OY 181 GlnAlaGlyAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGCGCGGGGCGCGCGAGGCGCGGAGCGGCGCTCAGGCGCATCCGAGCGGCGCTG 661
OY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGAGACAGGAGCGCGCTGCGGCGCGCCACTGTGGGCTCCCTGGCGGCGCAGCG 721
OY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgValArgMetGluGluMetGly 240
DB 722 CTACAGGAGCGGCGCCAGGCTGGGCGAGCGGCTGCGCGCGATGGAGAGATGGC 781
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCGAGACCGCGACCGCTGAGAGAGGTGAAGAGAGAGTGGCGGAGTGGCGGCAAG 841
OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGAGAGCGCCAGCAGATACCTGACAGCGGAGGCTTCCAGGCGGCGCTTCAG 901
OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGluLeuValGluLys 300
DB 902 AGCTGTGTCAGCGCTGCTGTCGAGACATGACAGCGGAGTGGCGGCGGCTGTGGAGAG 961
OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTCCAGAGCTGCGCTGGGACACACGCGCGCCCTGTGTCACAGCACAATCNC 1012
RESULT 14
ABL65450
ID ABL65450 standard; DNA; 1157 BP.
XX
AC ABL65450;
XX
XX 15-MAY-2002 (first entry)
DT
XX
XX Lung cancer related gene sequence SEQ ID NO:3787.
DE
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.

PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 03-OCT-2000; 2000US-237315P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237588P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -

PS Claim 1; SEQ ID 3787; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 3 44e-113 Length: 1157
 Score: 1561.00 Matches: 313
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.24% Indels: 0
 DB: 24 Gaps: 0

US-09-827-854-17 (1-317) x ABL65450 (1-1157)

QY 1 MetlyValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 Db 62 ATGAGGTTCTGTGGGTGGCTGTGCTGCACATTCCTGGCAGAGATCCAGGCCAAGGTG 121
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGluThr 40
 Db 122 GAGCAACGGCTGGAGACAGAGCGGAGCCGACCTGGCGCAGAGAACCGGTGGCAGAGC 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheThrAspTyrLeuArgTrpValGlnThr 60
 Db 182 GCCACGCGCTGGAGACTGGACTGGCTGCTTTGGGATTACCTGGCTGGTGGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnLeuLeuArgAla 80
 Db 242 CTGTCTGAGCAGGTCGAGAGAGAGTCGTGCTGACCTCCCAATCCACCCCAAGTGAAGGCG 301
 QY 81 LeuMetAspGluThrMetGlyGluLeuValAlaTyrIleSerGluLeuGluGluGlnLeu 100
 Db 302 CTGATGGACAGACACATGAAGGAGTTGAAGCCCTACAAATCGGAAGCTGGAGAAACATG 361
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerGlyGlnGlnAlaAlaGlnAla 120
 Db 362 ACCCGGCTAGCGAGAGAGACCGCGGACGCGCTCTCCAGAGAGCTGCAGAGCGCGAGCC 421
 QY 121 ArgLeuGluValAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
 Db 422 CGGCTGGCGCGGACATGAGAGAGAGTGTGGCGCGCTGCTGTCAGTACCGCGCGAGAGTG 481
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuValArgLeuAlaSerHisLeuArg 160
 Db 482 CAGGCCATGCTCGGCGAGACACCGAGAGAGCTCGGGTGGCTCCCTCCCTCCCTCCGCGC 541
 QY 161 LysLeuArgGlySerArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
 Db 542 AAGCTGGTAGAGCGGCTCTCCCGCATCCGATGACCTCGAGAACCGCCCTGCGCATGAC 601
 QY 181 GlnAlaGlyAlaArgGluGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 602 CAGCGCGGCGCGCGAGCGCGCGAGCGCGCTGACGCCCATCCGCGAGCGCGCTGGGG 661
 QY 201 ProLeuValGluGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 662 CCCCTGGTGAAACAGGCGCGCGTGGCGCGCCACGTGGCTCCCTGGCGCGCGCAGCGG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyArgLeuArgAlaArgMetGluMetGly 240
 Db 722 CTACAGAGAGGCGGCCAGAGCGCTGGCGAGCGCTGCGCGCGGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValGlyGluGlnValAlaGluValArgAlaVal 260
 Db 782 AGTCGGACCGCGAGCGCGCTGACGAGGTGAAGACAGAGTGGCGGAGGCGCGCGCAAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 842 CTGGAGAGAGAGGCCAGCAGATACCTGTCAGGCGCGAGCGCTTCAGGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnThrAlaGlyLeuValGluLys 300
 Db 902 AGCTGGTTCAGAGCCCTGTGTGAAGACATGACAGCGCATGAGGCGCGGCTGTGTGAAG 961
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 962 GTGACAGGTGCGTGGGACACAGCGCGCGCTGTGCTCCACGAGCAATATCAC 1012

RESULT 15
 AAN50450
 ID AAN50450 standard; DNA; 1110 BP.
 XX
 AC AAN50450;
 XX
 DT 09-JAN-1992 (first entry)


```

XX DE Sequence encoding human apolipoprotein E.
XX KW Hyperlipaemia; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 15..968
XX FT /*tag= a
XX FT mat_peptide 69..965
XX FT /*tag= b
XX PN JP60118189-A.
XX PD 25-JUN-1985.
XX PE 29-NOV-1983; 83JP-0224980.
XX PR 29-NOV-1983; 83JP-0224980.
XX PA (MITU ) MITSUBISHI CHEM IND KK.
XX DR WPI: 1985-188003/31.
XX DR P-PSDB; AAP51204.
XX PT DNA fragment - contg. DNA which codes human apolipoprotein E for
XX PT treatment of hyperlipaemia.
XX PS Claim 3; Page 484; 8pp; Japanese.
XX CC The sequence may be used to produce the apolipoprotein E, useful in the
XX CC treatment of hyperlipaemia.
XX SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;

Alignment Scores:
Pred. No.: 3,93e-113 Length: 1110
Score: 1560.00 Matches: 313
Percent Similarity: 98.74% Conservative: 0
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.17% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-17 (1-317) x AAN50450 (1-1110)
QY 1 MetIysValLeuTrpAlaIalaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 15 ATGAAGTTCTGTGGCTCTGCTGTCACATTCTTGGCAGATGCCAGGCCAAGGTG 74
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGACCCGAGCCGAGCTGCGCACAGACCGAGTGGCAGAGC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValaGlnThr 60
DB 135 GGCACACGCTGGGAACCTGACATGGGTCTGTTGGGATTACCTGGCGTGGGTCAGACA 194
QY 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTGACACCTGACAGAGAGCTGCACCTCCAGTCCACCCAGGAACCTGAGGGCG 254
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
DB 255 CTGATGGACGAGACCACTGAAGAGTTGAGGCCCTACAAATCGGAACCTGGAGAGACACTG 314
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaIaGlnAla 120
DB 315 ACCCGCGTGGCAGAGACGCGGACGCGCTGTCCAAAGAGCTGCAGCGCGCCAGGCC 374
QY 121 ArgLeuGlyAlaAspMetLysPvalCysGlyArgLeuValaGlnTyrArgGlyGluVal 140
DB 375 CGGCTGGCGCGGACATGAGGACGTGCGCGCTGTGTGACAGTACCGCGCGAGGTG 434

```

```

QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGGCCAGACACCGAGAGCTGCGGGTCTGCGCTCCACCTCGGCC 494
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 495 AAGCTGCGTAAACGGGCTCTCCGCGATGCGATGACCTCCAGAAAGCGCTGCGAGTGC 554
QY 181 GlnAlaGlyAlaArgGluGluValaGluArgGlyLeuSerValaThrArgGluLeuArgLeuGly 200
DB 555 CAGGCCGGGGCCCGGAGGGCCCGAGCGGCTCTAGCGCATCCGCGAGCGCTGGGG 614
QY 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGTGTGAACAGAGCGCGCGTGGCGCGCCACTGTGGCTCTCTGGCCGCGCACGCCG 674
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 675 CTACAGAGACGGGGCCAGGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGAGATGGGC 734
QY 241 SerArgThrArgAspArgLeuAspGluValaLysGluGlnValaIaGluValaArgAlaLys 260
DB 735 ACCCGGACCCGGGACCGCTGGACGAGGTGAAGAGACAGTGGCGAGGTGGCGCCGCAAG 794
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 795 CTGGAGAGACAGGCCCGACAGATACGCTGCGAGGCCGAGGCCCTTCCAGGCCCGCTCAAG 854
QY 281 SerTrpPheGluProLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValaGluLys 300
DB 855 ACCTGATTGAGCCCTGTGTGAAGACATGACAGCCGCACTGGGCGCGCTGTGTGAAGAG 914
QY 301 ValGlnAlaIaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTGCAAGCTGCCGTGGGACACAGCGCGCCCTGTGCCAGAGACAAATAC 965

```

Search completed: March 14, 2003, 12:43:15
 Job time : 183.728 secs

NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:

Pred. No.:	5,98e-101	Length:	1126
Score:	1123.00	Matches:	222
Percent Similarity:	83.12%	Conservative:	44
Best Local Similarity:	69.38%	Mismatches:	46
Query Match:	70.67%	Indels:	8
DB:	4	Gaps:	3

US-09-827-854-17 (1-317) x US-08-949-155-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 51 ATGAGGTTCTGTGGTTCTTTGGTGAACCTCTCGCAGATGCCGACAGAGGAC 110
|||||

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
|||
DB 111 GAGCCGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGGAGGCCAAG 158
|||

QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrp 57
|||||
DB 159 TGGCAGGAGCCAGCCCTGGGAGCAGCCCTGGGCCCTCTGGGATTACCTGGCCTGG 218
|||||

QY 58 ValGlnThrLeuSerGlnGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
|||||
DB 219 GTCCACTCCCTGTCTACCAAGTGCAGGAGGAGCTCTAGCACCAAGTCAACCCAGGA 278
|||||

QY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaThrLysSerGluLeuGlu 97
|||||
DB 279 CTGACGAGCTGATAGAGAGAGCATGAGAGGTGAGGCTACCGCAGAGAGCTGGAG 338
|||||

QY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
|||||
DB 339 GCGCAGCTGGCCCTGGCCAGCAGGAGCAGCCGCCCTGTGCTCAAGAGCTGGCAGGG 398
|||||

QY 118 AlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArg 137
|||||
DB 399 GCGCAGCCCGCTGGCCCGCAGCATGGAGAGCTGCCAACCCTTGGTCTCTACCCG 458
|||||

QY 138 GlyGluValGlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSer 157
|||||
DB 459 ACGAGGTGCACAACTGTTGGCCAGACCGAGAGCTGGGAGCGCCGTGGCTTCG 518
|||||

QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeu 177
|||||
DB 519 CACCTGCGCAAGCTGGCAGGCGGTCTCCGCGACACCGAGAGCTGCAGAAAGCCCTG 578
|||||

QY 178 AlaValTrpGlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGlu 197
|||||
DB 579 GCGGTGTACCAAGCGGGGCTGCCGAGGGCGCCGAGCGAGCTGACGCCCTCCCGGAG 638
|||||

QY 198 ArgLeuGlyProLeuValGluGlnArgValArgAlaAlaThrValGlySerLeuAla 217
|||||
DB 639 CGCTCGGGCCCTGTGTGAGCAGGGCCGATTGGCGCCGACCTGTGATACAGAGGCC 698
|||||

QY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 237
|||||
DB 699 GCGCAGCCCTCGCGCAGGCGCGGAAAGCTGGCGCCCAACCTGGCGGAGCGGCTGGAG 758
|||||

QY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValGlyGlnGlnValAlaGluVal 257
|||||
DB 759 GAGATGGAGCGGAGCCCGCAGCCCTGTGATGAGATCGTGAAGCAGCTGGAGAGAGTG 818
|||||

QY 258 ArgAlaLysLeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
|||||
DB 819 CCGACCAAGGTGAGAGCAGGAGCAGCAGCTTGGCCCTGACAGCCGAGGAGATTCCAGCCG 878
|||||

QY 278 ArgLeuLysSerTrpPheLeuProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 297
|||||
DB 879 CTCCTCAAAAGGCTGTTCAGGCTCTGTGTGGAAGACATACGCGCAGTGGCGCGGCTG 938
|||||

QY 298 ValGluLysValGlnAlaValGly---ThrSerAlaAlaProValProSerAspAsn 316
|||||
DB 939 GTGGAGAGATGACGTGGCGCGTGGAGCATTAAGCTCTCCACCTCTGCGCCAGTATAT 998
|||||

RESULT 3
US-09-819-964-5
Sequence 5, Application US/09819964
Patent No. 6369294
GENERAL INFORMATION:
Applicant: Piedrahita, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DUKKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 60/046,094
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:

Pred. No.:	5,98e-101	Length:	1126
Score:	1123.00	Matches:	222
Percent Similarity:	83.12%	Conservative:	44
Best Local Similarity:	69.38%	Mismatches:	46
Query Match:	70.67%	Indels:	8
DB:	4	Gaps:	3

US-09-827-854-17 (1-317) x US-09-819-964-5 (1-1126)

QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 51 ATGAGGTTCTGTGGTTCTTTGGTGAACCTCTCGCAGATGCCGACAGAGGAC 110
|||||

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
|||
DB 111 GAGCCGGG-----CCGCCCGGAGGTGCACGTGTGTGGAGGAGGCCAAG 158
|||

Oy	38	TpqlInSerIglYngIaNgTrpGlUleuAlaLeuLylarqPheTpaPyrIleuArTrp	57
Dd	159	TGGCAAGGCGACCAACCCCTGGGACAGAGCCCTGGGCGCCTTCTGGATTCACCTGGCTGG	218
Oy	58	ValGlnTrhIleuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrlngLnu	77
Dd	219	GTGCAGTCCCTGTCTGCACCAAGTCACAGAGAGAGCTGTCTACACCAAGATTCACCAGAA	278
Oy	78	LeuArqAlaLeuMetAspGluTrhMetIlySGLueuLySaIaTrYIlySSerGluLeuGlu	97
Dd	279	CTGACGGAGCTGATPAGAGAGATGAAGAGATGAAGGCTTAAGCCGAGAGACTGGAG	338
Oy	98	GluGlnLeuTrhProValAlaIglUlnThrPrrAlaIarLeuSerIySGLueGlnAla	117
Dd	339	GCGCAGCTGGGGCCCCGTGATACCAGAGACGAGGCGCGCTGTCCAAAGAGACTGCAGCG	398
Oy	118	AlaGlnAlaIarLeuGluYalaAspMetGluAspValYcIylarLeuValGlnTrYarq	137
Dd	399	GCGCAGGCGCCCGTGGGGCGCGCATGGAGAGACCTGGCGAACCCCTGTGCTTACCGC	458
Oy	138	GlyGluValGlnAlaMetLeuAspGlnSerThrlngLnuLeuArqValArqLeuAlaSer	157
Dd	459	AGCGAGGTGCACAAACATTTGGCCAGACCAACGAGAGAGCTGCGAGCCGCTGGCTTC	518
Oy	158	HisLeuArqIlySLeuArqIlySaIarLeuLeuArqAspAlaAspAspLeuGlnIlyScYleu	177
Dd	519	CACGTGGCCCAAGTGGCGCAACGGGCTGTCCGGACACACGAGAGACTGCAGAAAGCGCTG	578
Oy	178	AlaValIYrGlnAlaGlyAlaArgGluGlnYalaGluArqGlyLeuSerAlaIleArqGlu	197
Dd	579	GCCGTGTACCAAGCGGGGCTCGCGAGGGCCGCCAGCGACACTGCAGACGCTCTCCGAG	638
Oy	198	ArgLeuGlyProLeuValGlnGlnGlyArqValArqAlaIaIaThraValGlySerLeuAla	217
Dd	639	CGCCTCGGGCCCCCTGTGTGACAGAGGCCGATTCGGCGCGCCACCTCTGAGTACCAAGGCC	698
Oy	218	GlyGlnProLeuGlnGluArqAlaGlnAlaTrpGlyGluArqLeuArqAlaArqMetGlu	237
Dd	699	GGCCAGCGCGTGGCGAGCGCGCGGAAGCCTGGGGCCAGAACTGGCGCGAGCTGGAG	758
Oy	238	GluMetGlySerArqTrhArqAspArqLeuAspGluValIySGLuGlnValAlaGluVal	257
Dd	759	GAGATGGGCGACCCGAGCCCGCGACCGCTGGATGAAGTGGTGGAGCGAGGAGAGGTG	818
Oy	258	ArgAlaIySLeuGlnGluGlnAlaGlnGlnIleArqLeuGlnAlaGlnAlaIaPheGlnAla	277
Dd	819	CGCAACCAAAAGTGGAGAGACAGGAGCAAGTGGCGCTCGCAGCGCGGAGGATTCACAGCC	878
Oy	278	ArgLeuIySSerTrpPheGluProLeuValGluAspMetGlnArqGlnTrpAlaGlyLeu	297
Dd	879	CTCCCTCAAAAGCTGTGTTCAGACCTCTGTGTGAACACATACAGGCGCCAGTGGCGGCGCTG	938
Oy	298	ValGlyIySValGlnAlaIaIaValGly---ThrSerAlaIaIaProValProSerAspAsn	316
Dd	939	GTGGAGAGAGTGCAGTGGCGCGTGAACATTAAGCTCTCCACCTCTGCGCCCAAGATTAAT	998

RESULT 4
 US-08-949-155-51
 : Sequence 51, Application US/08949155
 : Patent No. 6271436
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Piedrahita, Jorge A
 : APPLICANT: Bazer, Fuller W
 :
 : TITLE OF INVENTION: Compositions and Methods for the
 : TITLE OF INVENTION: Generation of Transgenic Animal Species
 :
 : NUMBER OF SEQUENCES: 51
 :
 : CORRESPONDENCE ADDRESS:
 :

ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patent In Release #1.0, Version #1.30
6
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/08/949,155
9  FILING DATE: Concurrently Herewith
10 CLASSIFICATION: 800
11
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US 60/027,338
14 FILING DATE: 11-OCT-1996
15
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US 60/046,094
18 FILING DATE: 09-MAY-1997
19
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Hilder, David W.
22 REGISTRATION NUMBER: 41,071
23 REFERENCE/DOCKET NUMBER: TARK:177
24 TELECOMMUNICATION INFORMATION:
25 TELEPHONE: (512) 418-3000
26 TELEFAX: (713) 789-2679
27
28 INFORMATION FOR SEQ. ID NO.: 51:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 4267 base pairs
31 TYPE: nucleic acid
32 STRANDEDNESS: single
33 TOPOLOGY: linear
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
```

Alignment Scores:	
Pred. No.:	6, 99e-87
Score:	987.00
Percent Similarity:	57.988
Best Local Similarity:	48.548
Query Match:	62.11%
DB:	4
	Gaps:
	5

Qy	1	MetylsValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	2448	ATGGCTGTGAACGTGGCTCGGT-----GCAGATGCCCGACACAGAGAC	2485
Qy	21	GIuGlnAlaValGIuThrGIuProGIuProGIuLeuArg-----GlnGlnThrGlu	37
Db	2490	GAGCGCGGAG-----CGCGCGCGAGAGTGCACGTGTGGAGGAGACCCACAG	2537
Qy	38	TrpGlnSerGIuArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrp	57
Db	2538	TGGCAGGGCAGCCAGCCCTGGAGACAGCCCTGGCGGCTTGTGGATTACTGGCTGG	2597
Qy	58	ValGIuThrLeuSerGIuGlnValGIuGlnGluLeuLeuSerSerGlnValThrGlnGlu	77
Db	2598	GTCGAGTCCCTGTCTGTGACCAAGTGACAGAGAGACTGCTCAGCACCAAGGTACACCCAGGAA	2657
Qy	78	Leu-----	78
Db	2658	CT-GACGTAAAGTCCACCCGACACTCCCGCGCGCGCGCGCGCGCGCGCGCTGA	2716
Qy	78	-----	78
Db	2717	CCCTCTGGCGAACCGTGTGTCTGGACCTCAGGCTCCACCCGTCGGGTTTCCTCTCG	2776
Qy	78	-----	78
Db	2777	TCTCTGTGCGCAACTCTGGGGGTCTGGGTCTCGTTTCCTTTTTCCTTCCTTCCTTTT	2836
Qy	78	-----	78
Db	2837	TTGGGGGGAAAAAACTTTTCTTTCTTTTCTTTTCATTTCATTGACTTCATGCTTGGCTTCCTTCC	2896
Qy	78	-----	78

Db	2837	TTGGGGGGAAAAA	CTTTTCTTTTCTTTTCTTTTCTTTGACTTGGACTCATGCTGCTTTCTTCC	28936
QY	78	-----	-----	78
Db	2897	ATCTTGACCTCC	CTTCCTTCGCTCTCTGGGTCAGTCTTGGCGTCCCTTCGTCTCTCG	29566
QY	78	-----	-----	78
Db	2957	AATCTCTGGCAC	GTCTGGCCATCCGACGTACAGGACCCCTCTTCTCCCTCACGCC	30166
QY	79	-----	-----	92
Db	3017	CCCGCCCTCTC	TGCGCCACAGGAGCTGATAGAGAGACATGAGAGGTGAAGGCTTAC	30766
QY	93	LysSerGluLeu	GluGluGluLeuThrProValAlaGluGluThrArgAlaArgLeuSer	112
Db	3077	CGCGAGGAGT	CGAGGCGAGCTGGGCCCCCTGACCCAGAGAACACAGCGCGCTCTGCC	31366
QY	113	LysGluLeuGln	AlaAlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArg	132
Db	3137	AAGGAGCTGC	AGGCGGCGGACGCCCGCGTGGCGCGCATGAGAGACCTGCGCAACGC	31966
QY	133	LeuValGlnTyr	ArgArgGlyValAlaAlaMetLeuAspGlnSerThrGluLeuArg	152
Db	3197	TTGGTGCTCT	ACCCGACAGGAGTGCACACATTTTGGCGACACCCAGAGAGCTCGGG	32566
QY	153	ValArgLeuAla	SerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp	172
Db	3257	AGCGCCCTGT	CGCTCCACCTGCCACAGCTGGCCACACCGGCTCTCGCGACACCGAGAC	33166
QY	173	LeuGlnTyrCys	LeuAlaValTyrGlnAlaGlyAlaArgGluGlyValGluArgGlyLeu	192
Db	3317	CTGCAGAGCC	CTCGGCGGTCTACAGGCGGGCTGCGGAGGGCGCGAGCGCAGCGTG	33766
QY	193	SerAlaIleArg	GluArgLeuGlyProLeuValGluGlnGlyArgValArgAlaIleThr	212
Db	3377	AGCGCCCTCG	CGGAGCGGCTCGGGCCCCCTGGTGGACAGGGCGCATGCGCGCCGAC	34366
QY	213	ValGlySerLeu	AlaGlyGlnProLeuGlnGluArgAlaGlnAlaTyrPylGluArgLeu	232
Db	3437	CTGAGTACCA	GGCGGCGACGCGCTGCGGACGCGCGGACCGCGGAGACCTGGGGCGAGAACGTG	34966
QY	233	ArgAlaArgMet	GluGluMetGlySerArgThrArgAspArgLeuAspGluValLysGlu	252
Db	3497	CGCGGAGCGG	CTGGAGGATGGGACGCGGACCGCGGACCGGCTGTGATGATGCTGAG	35566
QY	253	GlnValAlaGlu	ValAlaArgAlaLysLeuGluGlnAlaGlnAlaIleArgLeuGlnAla	272
Db	3557	CAGCTGAGAG	AGAGTGGCCACCAAGTGGAGAGACGACGACGACCATGTCGCTGAGGCC	36166
QY	273	GluAlaPheGln	AlaArgLeuLysSerTyrPheGluProLeuValGluAspMetGlnArg	292
Db	3617	GAGGAGATTCC	ACGCGCTCCCAAAAGCTGTTCGACCTCTGTGGAAACATACGCGCG	36766
QY	293	GlnThrPheGly	LeuValGluLysValGlnAlaAlaValGly---ThrSerAlaIlePro	311
Db	3677	CAGTGGCGGG	CGTGGTGGAGAGATGACAGTGGCGGTGAGCATAGCTCTCCACCTCT	37366
QY	312	ValProSerAsp	asn 316	
Db	3737	GCGCCCACTG	ATAT 3751	
RESULT 6				
US-08-726-306A-28				
: Sequence 28, Application US/08726306A				
: Patent No. 5958684				
: GENERAL INFORMATION:				
: APPLICANT: van Leeuwen, Frederik Willem				
: APPLICANT: Burbach, Johannes Peter Henri				
: APPLICANT: Grosveld, Franklin G.				
: TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS				
: NUMBER OF SEQUENCES: 189				
: CORRESPONDENCE ADDRESS:				

	ADDRESSEE: Banner & Wilcoff, Ltd.
	STREET: 1 Financial Center
	CITY: Boston
	STATE: MA
	COUNTRY: US
	ZIP: 02111
	COMPUTER READABLE FORM:
	MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
	COMPUTER: IBM PC compatible
	OPERATING SYSTEM: PC-DOS/MS-DOS
	SOFTWARE: WordPerfect 6.1
	CURRENT APPLICATION DATA:
	APPLICATION NUMBER: US/08/726,306A
	FILING DATE: 02-Oct-1996
	PRIOR APPLICATION DATA:
	APPLICATION NUMBER: GB 95/20080.4
	FILING DATE: 02-Oct-1995
	PRIOR APPLICATION DATA:
	APPLICATION NUMBER: US 60/009,832
	FILING DATE: 01-Jan-1996
	ATTORNEY/AGENT INFORMATION:
	NAME: Williams, Ph.D., Kathleen M.
	REGISTRATION NUMBER: 34,380
	REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
	TELECOMMUNICATION INFORMATION:
	TELEPHONE: (617) 345-9100
	TELEFAX: (617) 345-9111
	INFORMATION FOR SEQ ID NO: 28:
	SEQUENCE CHARACTERISTICS:
	LENGTH: 660 base pairs
	TYPE: nucleic acid
	STRANDEDNESS: double
	TOPOLOGY: linear
	MOLECULE TYPE: cDNA
	US-08-726-306A-28
	Alignment Scores:
	Pred. No.: 3.23e-86 Length: 660
	Score: 969.00 Matches: 195
	Percent Similarity: 97.99% Conservative: 4
	Best Local Similarity: 97.99% Mismatches: 0
	Query Match: 60.98% Indels: 0
	DB: 2 gaps: 0
US-09-827-854-17 (1-317) x US-08-726-306A-28 (1-660)	
Oy 1 MetLysValLeuTrpPAlaAlaLeuLeuValThrPhleuAlaIagLYcSGlnAlaIysVal 20	
Dd 62 ATGAAGGTTCTGTGGGCGTTCGTGCTGCATTCCTCGCAGAGATGCCAGGCCAAAGTG 122	
Oy 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40	
Dd 122 GAGCAAGCGGTGGAGACAGACC66GAGCCGAGCTGCGCCAGCACAGACCGAGTGGCAAGC 181	
Oy 41 GlyGlnArgTPGLuLeuAlaIaLeuGlyArgPheTrpAspyrIleuArgTPVAlaGlnThr 60	
Dd 182 GCCCGCGCTGGGAACGTGGCACTGGGTGCTTTGGGATTACTGGCGCTGGGTGCAGACA 241	
Oy 61 LeuSerGluGlnValaGlnGluGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80	
Dd 242 CTGTGTGAGCGAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTCAACCCAGAAGCATGAGGCG 301	
Oy 81 LeuMetAspGluThrMetLysGluLeuLeuAlaTyrllysSerGluLeuGlnGluGlnLeu 100	
Dd 302 CTGATGAGACGAGACCATTAAGAGATTGAAGCCCTTCAAATCGGAACATGGAGAACACTG 361	
Oy 101 ThrProValaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120	
Dd 362 ACCCGGTAGCGGAGAAACACGCGGACGCGCTGTCCAAGGAGCTGACAGCGCGCAGGCC 421	
Oy 121 ArgLeuGlyAlaAspMetGluAspValaCylserGlyArgLeuValaGlnTyrrArgGlyGluVal 140	
Dd 422 CGGCTGGGCGCGGACATGGAGAGAGCTGTGGCGCGCTGTGTGCAGTACCGCGCGAGAGTG 481	

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:

Pred. No.:	5,73e-28	Length:	252
Score:	367.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best Local Similarity:	96.43%	Mismatches:	3
Query Match:	23.10%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-17 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAGAGACTGCAGCGCCGCCAGCCGGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GTGTGGCGC-CGCTGTGTGACGTACCGCGGAGGTGACAGCCATGCTCGGCCAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGACTGTGGGTGGCTGCGCTCCACCTCGCAGCTGCGTAAGCGGCTCTCCCGC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTGCAGAACTCCTGCGCAGTGTACACAGCCGGGCGCGGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-567-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:

Pred. No.:	5,73e-28	Length:	252
Score:	367.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best Local Similarity:	96.43%	Mismatches:	3
Query Match:	23.10%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-17 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerIysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAGAGACTGCAGCGCGGCCAGCCGGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuAspGlnSerThr 148
|||||
DB 62 GTGTGGCGC-CGCTGTGTGACGTACCGCGGAGGTGACAGCCATGCTCGGCCAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
|||||
DB 121 GAGGACTGTGGGTGGCTGCGCTCCACCTCGCAGCTGCGTAAGCGGCTCTCCCGC 180

QY 169 AspAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACTGCAGAACTCCTGCGCAGTGTACACAGCCGGGCGCGGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 5,73e-28
Score: 367.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.10%
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-682-24 (1-252)

QY 109 AlAargLeuserLysGluLeuGlnAlaGlnAlaAargLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGGCGCGCCGCTGGCGCGGACATGGAGAGAC 61
QY 129 ValCysGlyAargLeuValGlnTyrAargGlyGluValGlnAlaMetLeuAspGlnSerThr 148
Db 62 GTGTGGCC-CGCTGTGGAGTACCGCGGAGGTGCAAGCCATGCTGGCCAGAGCACC 120
QY 149 GluGluLeuAargValAargLeuAlaSerHisLeuAargLysLeuAargLysAargLeuAarg 168
Db 121 GAGAGAGCTGGGGGTGGCGCTCCCTCCACTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspaAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaAargGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGCGAGTACCAAGCCGCGGCGCGCGAGGGCGCC 240
QY 189 GluAargGlyLeu 192
Db 241 GAGCGGGGCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287, 679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 5,73e-28
Score: 367.00
Percent Similarity: 96.43%
Best Local Similarity: 96.43%
Query Match: 23.10%
Matches: 81
Conservative: 0
Mismatches: 3
Indels: 1
Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-679-24 (1-252)

QY 109 AlAargLeuserLysGluLeuGlnAlaGlnAlaAargLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGGCGCGCCGCTGGCGCGGACATGGAGAGAC 61
QY 129 ValCysGlyAargLeuValGlnTyrAargGlyGluValGlnAlaMetLeuAspGlnSerThr 148
Db 62 GTGTGGCC-CGCTGTGGAGTACCGCGGAGGTGCAAGCCATGCTGGCCAGAGCACC 120
QY 149 GluGluLeuAargValAargLeuAlaSerHisLeuAargLysLeuAargLysAargLeuAarg 168
Db 121 GAGAGAGCTGGGGGTGGCGCTCCCTCCACTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspaAlaAspAspLeuGlnLysCysLeuAlaValTyrGlnAlaGlyAlaAargGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGCGAGTACCAAGCCGCGGCGCGAGGGCGCC 240
QY 189 GluAargGlyLeu 192
Db 241 GAGCGGGGCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 5,73e-28 Length: 252
Score: 367.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.10% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-17 (1-317) x US-09-397-766-24 (1-252)

QY 109 AAlaAglleuSerlyGluLeuInlaAlaGlnAlaAryLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGGGCGGCGGCTGGGCGGACATGGAGGAC 61
QY 129 VALCYSGLYARGLLeuValGlnTYrArGlyGluValAlGlnAlaMetLeuAspGlnSerThr 148
DB 62 GGTGTGGC-CGGCTGTGTGAGTACCGCGCGGAGGTGCAGGCCATGCTCGCCAGACACC 120
QY 149 GUGLULeuArValArGLeuAlaSerHisLeuArGlySLeuArGlyArGLeuLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTCCCTCCACCTGCGCAAGCTCGTAAAGCGGCTCTCCGC 180
QY 169 AAPAlaAspAspLeuGlnLysCysLeuAlaValTYrGlnAlaGlyAlaArGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGGCAGTGTACAGCGCGGCGCGCGAGGGCGCC 240
QY 189 GUArGlyLeu 192
DB 241 GAGCGGGGCTC 252

RESULT 15
US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 627573
GENERAL INFORMATION:

APPLICANT: K Ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 5,73e-28 Length: 252
Score: 367.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.10% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-17 (1-317) x US-09-287-681-24 (1-252)

QY 109 AAlaAglleuSerlyGluLeuInlaAlaGlnAlaAryLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGGGCGGCGGCTGGGCGGACATGGAGGAC 61
QY 129 VALCYSGLYARGLLeuValGlnTYrArGlyGluValAlGlnAlaMetLeuAspGlnSerThr 148
DB 62 GGTGTGGC-CGGCTGTGTGAGTACCGCGCGGAGGTGCAGGCCATGCTCGCCAGACACC 120
QY 149 GUGLULeuArValArGLeuAlaSerHisLeuArGlySLeuArGlyArGLeuLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTCCCTCCACCTGCGCAAGCTCGTAAAGCGGCTCTCCGC 180
QY 169 AAPAlaAspAspLeuGlnLysCysLeuAlaValTYrGlnAlaGlyAlaArGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGGCAGTGTACAGCGCGGCGCGCGAGGGCGCC 240
QY 189 GUArGlyLeu 192
DB 241 GAGCGGGGCTC 252

Search completed: March 14, 2003, 20:18:46
Job time : 40.7343 secs


```

Db      661 CCCCTGGTGGAAAGGCGCGCTGCGGCCGACACTGTGGCTCCCTGGCCGCGCAGCCG 720
Qy      221 LeuGluGluArgAlaGlnAlaIleThrPgluArgLeuArgAlaIleThrPgluGluGly 240
Db      721 CTACAGAGCGGCGCCAGGCGCTGGGCGAGCGCGCTGGCGCGGATGGAGATGGGC 780
Qy      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db      781 AGCCGAGCCCGGACCGCTGGAGCGAGTGGAAGGAGCAGTGGCGAGGTGGCGCGCAG 840
Qy      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db      841 CTGGAGGAGCAGCGCCAGCAGATAGCCCTGCGAGCGCGAGCCCTCCAGCGCCCTCAAG 900
Qy      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db      901 ACCTGCTTCGACCCCTGGTGGAGACATGACAGCGCAATGGGCGCGGTGGTGGAGAG 960
Qy      301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db      961 GTGCAAGCTGCGCGTGGGCGACAGCGCGCCCTGTGTCCAGCAGACATATAC 1011

```

RESULT 3

```

; US-09-870-759-129
; Sequence 129, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION:
US-09-870-759-129

```

Alignment Scores:

```

Pred. No.: 1,05e-135      Length: 1156
Score: 1570.00           Matches: 315
Percent Similarity: 99.37%      Conservative: 0
Best Local Similarity: 99.37%      Mismatches: 2
Query Match: 98.80%           Indels: 0
DB: 9                      Gaps: 0

```

US-09-827-854-17 (1-317) x US-09-870-759-129 (1-1156)

```

Qy      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db      61 ATGAAGTTCTGTGGCTGCGTGTGTCACATTCTGTGGAGATGCCAGGCGCAAGGTG 120
Qy      21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
Db      121 GAGCAAGCGGTGGAGACAGAGCGGAGCGCGAGCTGCGCACACAGACAGAGGCGCAGAGC 180
Qy      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db      181 GCCCAGCGCTGGAGACTGCACACTGGCTTTGGGATTACTCGCGCTGCGCAGAGACA 240
Qy      61 LeuSerGluGlnValGluGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db      241 CTGTCTGACAGGTGAGAGAGAGCTGCTCAGCTCCAGAGTACCACCGAGAACTGAGGGCG 300
Qy      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100

```

```

Db      301 CTGATGAGACAGACCATGAAAGAGTTGAAGGCGCTCAAAATCGCAACTGGAGAGAACACTG 360
Qy      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db      361 ACCCGGTGGCGAGAGAGCGGGGACGCGTGTCCAAAGACTCCAGCGCGCGCAGAGGC 420
Qy      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTyrArgGlyGluVal 140
Db      421 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGTGTACATACCGCGCGAGGTG 480
Qy      141 GlnAlaMetLeuAspGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db      481 CAGCGCATCTCGCGCAGAGACACCGAGAGCTGCGGGTGGCGCTGCCCTCCACCTGCGC 540
Qy      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
Db      541 AAGCTGCGTAAACGGCGCTCCCGCATGCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 600
Qy      181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgLysArgLeuGly 200
Db      601 CAGCGCGGGGCGCGGAGGCGCGGCGCTCAGCGCATCCGAGCGCGCGCTGGGG 660
Qy      201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db      661 CCCCTGTGTGAACAGAGCGCGCTGCGGCGCGCACCTGTGGCTCCCTGGCGCGCAGCGG 720
Qy      221 LeuGlnGluArgAlaGlnAlaIleThrPgluArgLeuArgAlaIleThrPgluGluGly 240
Db      721 CTACAGAGCGGCGCCAGGCGCTGGGCGAGCGCGCTGGCGCGCGATGAGAGATGGGC 780
Qy      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db      781 ACCCGAGCCCGGACCGCTGGAGCAGAGTGAAGAGACAGTGGCGAGGTGGCGCGCAG 840
Qy      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db      841 CTGGAGAGCAGCGCCAGCAGATAGCCCTGCGAGCGCGAGCGCTTCAGAGCGCGCGCTCAAG 900
Qy      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db      901 ACCTGCTTCGACCCCTGGTGGAGACATGACAGCGCCCAATGGGCGCGGTGGTGGAGAG 960
Qy      301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db      961 GTGCAAGCTGCGCGTGGGCGACAGCGCGCCCTGTGTCCAGCAGACATATAC 1011

```

RESULT 4

```

; US-09-802-640-17
; Sequence 17, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Kleyen Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
; OTHER INFORMATION: (APOE)
US-09-802-640-17

```



```
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2244
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020124981A1 M12529
US-09-880-107-2244
```

Alignment Scores:

Pred. No.:	7.03e-135	Length:	1157
Score:	1561.00	Matches:	313
Percent Similarity:	98.74%	Conservative:	0
Best Local Similarity:	98.74%	Mismatches:	4
Query Match:	98.24%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-17 (1-317) x US-09-880-107-2244 (1-1157)

```
QY 1 MettysValleutrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 62 ATGAGGTTCTGTGGCTGGTGTCTGTGCATCTCTGGCAGAGATCCAGGCAAGTG 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
    |||||||
DB 122 GAGCAACGGGTGGAGACAGCCGAGCCGAGCTCGCCAGACAGACCGAGTGCGACAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||||
DB 182 GGGCAGGCGTGGAGACTGGACTGGCTGGCTTTGGGATTACCTGGCTGGTGCCAGACA 241
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
DB 242 CTGCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCCAATCCCAAGAACTAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
    |||||||
DB 302 CTGATGACAGACACCATGAAGAGTTGTAAGCCTACAAATCGGAACCTGGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
    |||||||
DB 362 ACCCGGTTAGCGAGAGAGACCGGGCCAGGCTGTCCAAAGAGCTGCCAGACGGCCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
    |||||||
DB 422 CGGCTGGCCCGCAGATGAGAGAGCTGTGGCGCGCTGTGTCAGTACCGCGGAGAGTG 481
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
DB 482 CAGGCGCATGCTGGCCAGACACCGAGAGCTGGGGTGGCTCGCTCCCAACCTGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
    |||||||
DB 542 AAGCTGCGTAAGCGGCTCTCCCGGATCCCGATGACTCGAGAGCGCCTGGCGAGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGluAlaGluArgGlyLeuSerAlaIleArgGluTrpGlnGly 200
    |||||||
DB 602 CAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGGCGCATCCGCGAGCGCTGGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
```

```
|||||
DB 662 CCCCTGTGGACAGGGCCCGGCGGCCAGCACTGTGGCTCCCGGCGCCAGACCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
    |||||||
DB 722 CTACAGAGCGGGGCCAGGCTGTGGCGAGCCGCTGCGCGCGGATGAGAGATGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
    |||||||
DB 782 AGTGGACCGCGCAGCCCTGTGAGAGAGGTGAGAGAGAGTGGCGGAGGTGGCGCCAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
    |||||||
DB 842 CTGGAGAGCAGGCGCCAGCAGATACGCTCAGGCGGAGGCTTCCAGGCGCCCTCAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGluArgGlnTrpAlaGlyLeuValGluLys 300
    |||||||
DB 902 AGCTGTTGAGACCCCTGTGTGAGAGACATCGACGCCAGTGGCGGCTGTGTGAGAG 961
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||||
DB 962 GTGAGGCTGCGTGGGCGACAGGCGCGCCCTGTGCGCCAGCAGCAATCAC 1012
```

RESULT 11

```
US-09-827-854-7
; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7
```

Alignment Scores:

Pred. No.:	1.32e-134	Length:	1156
Score:	1558.00 <td>Matches:</td> <td>314</td>	Matches:	314
Percent Similarity:	99.05% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.05% <td>Mismatches:</td> <td>3</td>	Mismatches:	3
Query Match:	98.05% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-17 (1-317) x US-09-827-854-7 (1-1156)

```
QY 1 MettysValleutrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
DB 61 ATGAGGTTCTGTGGCTGGTGTCTGTGCATCTCTGGCAGAGATCCAGGCAAGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
    |||||||
DB 121 GAGCAACGGGTGGAGACAGCCGAGCCGAGCTGCGCCAGACAGACCGAGTGCGAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||||
DB 181 GGGCAGGCGTGGAGACTGGAGTGGCTTTGGGATTACCTGGCGCTGGTGCGAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
DB 241 CTGCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACCTAGAGGGCG 300
```


Oy	158	stlenurtrgylsleuaatgylsarsrleuileumrgrspblaaaspaaleuglnltyscysleuall	178
Db	21336	CCTGGCGAAGCTGyGTAAAGCGGCTCTCCGGCATGGCTACCTGCAAGAGyGCTGGC	21395
Oy	178	aValTrglnalagialyalaaraaglucllyalagluaragsglyleuSerlaailaargluar	198
Db	21396	AGGTACCAAGAGCGCGGGGCGCGGAGGGCGCCGAGCGGGCTCTACAGCGCATCCGAGAGCG	21455
Oy	198	gleucllyProleuValgluglnGlyargValaarglaaIamprValgllySerleuAlaGl	218
Db	21456	CCTGGGGCCCCGTGGGGAACAGAGGCGCGTGGGGCGCATGTGGGTCTCCCTGGCGGG	21515
Oy	218	ylGlnProleuGlnGlnuargAlaGlnalAtrpGlygluarGleuAargAlaIarMetclugl	238
Db	21516	CCAGCGCGTACAGAGAGCGGGGCCCGAGGCGGCTGGGGCGAGCGGCTGGCGCGCATGGAGA	21575
Oy	238	uMetclSerarGthrArgrasparGleuaspGluVallysgluGlnValAlaGluValAr	258
Db	21576	GATGGGCAAGCCGAGACCCGCGACCGCTGGATCCAGGTGAAGGAGCAAGGTGGCGAGGTGGC	21635
Oy	258	gAlaIysleuGlnugluGlnAlaGlnGlnIleIarGrgleuGlnAlaGlnAlaIaphGlnAlaAr	278
Db	21636	CGCCAAAGCTGGAGGAGCAAGGCCCAAGACATACGCTTCGAGAGGCCGAGGCTTCCAGGCCCCG	21695
Oy	278	gleuLysSerTrpPheGlnProleuValGlnaspMetGlnarGlnTrpAlaGlyleuVa	298
Db	21696	CCTCAAGAGCTGTTGCGAGCCCCCTGGTGGAACATCGACGCCCAAGTGGGGCGGGGTGCT	21755
Oy	298	lGluIysValGlnAlaAlaValaGlyTrpSerAlaAlaProValProSerAspAsmHis	317
Db	21756	GGAGAAAGGTGACGAGCTGGCCGTGGGACACAGCCCGCCCTGTGGCCACGACGAATATCAC	21813

```

RESULT 13
US-09-925-302-133
: Sequence 133: Application US/09925302
: Patent No. US20020004941A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA104
: CURRENT APPLICATION NUMBER: US/09/925, 302
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05918
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124, 270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 133
: LENGTH: 786
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-302-133

```

Alignment Scores:	
Pred. No.:	5.06e-70
Score:	852.50
Percent Similarity:	97.31%
Best Local Similarity:	97.31%
Query Match:	53.65%
DB:	10
Length:	786
Matches:	181
Conservative:	0
Mismatches:	5
Indels:	2
Gaps:	0

US-09-827-854-17 (1-317) x US-09-925-302-133 (1-786)

Oy 132 ArgueuValGlnTyrArgGlyGlnValGlnAlaMetLeuaspGlnSerThrGlutIleu 151
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 CGCGTGTCAGTAGCCGGCGAGGTGCASGCCATGCTGGCCAGAGACCGAGAGCTG 65

Db 6 CGCCTGGTGCAGTACCGCGGCGAGGTGCASGCCATGCTCGGCCAGAGCACCAGAGGCTG 65

Oy 152 ArqValArgLeuAlaSerHisLeuArgCysLysLeuArgLysArgLeuAlaGlyAspalaasp 171
|||||
Db 66 CGGGTGCCTCGCCCTCCACACTGCCAA-CTCCGTAA-CGGCTTCCTCCGGCATGCCGAT 123

QY	172	AspleuglnIlyscysLeuAlaValTyrGlnAlaIleArgIleValArgIleValArgIly	191
Db	124	GACCTGCAGAAAGCGCTGTGCAGGTACACAGGCCGGGGCCCGCAGAGGCCGCAGACGGCC	183
QY	192	LeuSerAlaAlaIleArgGluArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAla	211
Db	184	CTCACCGCCATCCCGAGACGCCCTGGGGCCCTCGTGTGAACAGGGCCCGCTGGCGGGCCGC	243
QY	212	ThrValGlySerLeuAlaGlyAlaProLeuGlnGluArgAlaGlnAlaTyrPglGluArg	231
Db	244	ACTGTGGGCTCCCTGTGGCCGGCCAGCCGCTACAGAGAGCGGGCCAGCGCTGTGGCGAGCGG	303
QY	232	LeuArgAlaArgMetGluGluMetGlySerArgThrArgAspArgLeuAspGluValLys	251
Db	304	CTGGCCGGCGGATGGAGAGATGGGACGCCGACCCGCAACCCCTGTGACGAGAGGTGAAG	363
QY	252	GluGlnValAlaGluValArgAlaLysLeuGlnGlnGlnGlnAlaGlnGlnIleArgLeuGln	271
Db	364	GAGCAGGGTGGGAGGTGCTGGCCCAACTGGAGGAGGACGCCACGATACGCTTGGCAG	423
QY	272	AlaGluAlaPheGlnAlaArgLeuLysSerTrpPheGluProLeuValGluAspMetGln	291
Db	424	GCCGAGGCGCTTCCAGAGCGCCGCTCAAGAGCGTGTTCAGAGCCCGCTGCGGAAAGCATGGCAG	483
QY	292	ArgGlnTrpAlaGlyLeuValGluLysValGlnAlaAlaValGlyThrSerAlaAlaPro	311
Db	484	CGCCAGTGTGGCGGGGTGGTGTGAAGAGGTGCAGGCTCCGTGGGACACACGCCGCCCTT	543
QY	312	ValProSerAspAsnHis	317
Db	544	GTGGCCAGCGACATTCAC	561

```

RESULT 14
US-09-964-824A-374/c
; Sequence 374, Application US/09964824A
; Patent No. US20020102531A1
;
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; TITLE OR INVENTION: Sets
; FILE REFERENCE: 688290-73
; CURRENT APPLICATION NUMBER: US/09/964, 824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 374
; LENGTH: 478
;
; TYPE: DNA
; ORGANISM: Homo sapiens
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(478)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

```

Alignment Scores:	
Pred. No.:	2,16e-49
Score:	625.000
Percent Similarity:	97.69%
Best Local Similarity:	96.92%
Query Match:	39,33%
DB:	10
Length:	478
Matches:	166
Conservative:	1
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-854-17 (1-317) x US-09-964-824A-374 (1-478)

QY 1 MetysValLeuTrpPalaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

Db 396 ATGAGGTCTGTGGGCTGGCTGCTGCTCACAATTCCTGGCAGATGCCAGCCAAAGGTG 337
QY 21 GtuglnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCACTGGGCCAGACAGCCAGTGGCAGAGC 277
QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCTGGCAGACA 217
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTGTCTGAGAGGTGACAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAGGAGGCG 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGCTTGAAGGCTTCAAAATCGGAACGTGAGAACACT 97
QY 100 uThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGGCGGCGGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9
RESULT 15
US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; NAME/KEY: unsure
; LOCATION: (1)-(478)
; OTHER INFORMATION: n = a o r c o r g o r t
US-09-880-107-2491
Alignment Scores:
Pred. No.: 2,16e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.698 Conservative: 1
Best Local Similarity: 96.928 Mismatches: 2
Query Match: 39.338 Indels: 1
DB: 10 Gaps: 0
US-09-827-854-17 (1-317) x US-09-880-107-2491 (1-478)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 396 ATGAAAGTTCTGTGGGCTCGTTGCTGACATTCCTGGCAGGATGCCAGGCAAGGTG 337
QY 21 GtuglnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40
|||||
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCACTGGGCCAGACAGCCAGTGGCAGAGC 277
QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCTGGCAGACA 217
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTGTCTGAGAGGTGACAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAGGAGGCG 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGCTTGAAGGCTTCAAAATCGGAACGTGAGAACACT 97
QY 100 uThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGGCGGCGGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9

Search completed: March 15, 2003, 03:04:37
Job time : 82.8753 secs

QY 41 GtGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 276 GGGCAGCGCTGGAGACTGGACGTGGCTTGGATTACCTGGCTGGCTGGCTGGCAGACA 217
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
|||||
Db 216 CTGTCTGAGAGGTGACAGAGGAGCTGCTGACCTCCAGCTCCAGCAGGAGGAGGCG 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
|||||
Db 156 GCTGATGAGACAGACCATGAGAGGCTTGAAGGCTTCAAAATCGGAACGTGAGAACACT 97
QY 100 uThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
|||||
Db 96 GACCCGCGTGGCGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGGCGGCGGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
|||||
Db 36 CCGGCTGGCGCGACATGAGGAGCGTG 9

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-17
Perfect score: 1589
Sequence: 1 MKVLMALVTFLLAGCOAKV.....VEKVOAAVGTSAAPVSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09827854/runat.11032003.101610.27495/app.query.fasta.1.3576
-DB=EST -QFMT=fastp -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-ONITS-bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=5001 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.ecgn.1.1.8856@runat.11032003.101610.27495 -NCPU=6 -ICPU=3
-NO_XILPY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :
EST : *
1: em_estba : *
2: em_estchm : *
3: em_estlin : *
4: em_estlmu : *
5: em_estlov : *
6: em_estlpl : *
7: em_estlro : *
8: em_hlc : *
9: gb_estl1 : *
10: gb_estl2 : *
11: gb_hlc : *
12: gb_est3 : *
13: gb_est4 : *
14: gb_est5 : *
15: em_estfun : *
16: em_estom : *
17: gb_gss : *
18: em_gss_hum : *
19: em_gss_inv : *
20: em_gss_dln : *
21: em_gss_vrt : *
22: em_gss_fun : *
23: em_gss_mam : *
24: em_gss_mus : *
25: em_gss_other : *
26: em_gss_pro : *
27: em_gss_rtd : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1275.5	80.3	1027	13	B1670367	B1670367 603292738
2	1239	78.0	842	12	B6763371	B6763371 602735433
3	1233	77.6	933	13	B1668318	B1668318 603295681
4	1205	75.8	938	12	B6761746	B6761746 602717942
5	1202.5	75.7	942	13	B1600906	B1600906 603249241
6	1188	74.8	942	13	BM042094	BM042094 603615713
7	1187.5	74.7	927	12	B6472299	B6472299 602513830
8	1184.5	74.5	922	13	B1597743	B1597743 603248609
9	1175.5	74.0	817	12	B6774871	B6774871 602649975
10	1174	73.9	811	13	B1600563	B1600563 603244936
11	1160.5	73.0	845	12	B6829472	B6829472 602763768
12	1137	71.6	790	12	B6707147	B6707147 602670283
13	1134	71.4	919	13	B1551475	B1551475 603194314
14	1124	70.7	757	13	BM042228	BM042228 603616186
15	1114.5	70.1	907	12	B6706129	B6706129 602669093
16	1113	70.0	706	14	BM728696	BM728696 UI-E-EJ0-
17	1113	70.0	741	12	B6762924	B6762924 602735153
18	1111	69.9	741	11	AK010261	AK010261 Mus. muscu
19	1109	69.8	804	12	B6707252	B6707252 602684616
20	1108	69.7	834	13	BM042676	BM042676 603616028
21	1105.5	69.6	808	13	B1668329	B1668329 603295692
22	1102	69.4	855	13	B161362	B161362 602865769
23	1099	69.2	803	13	B1670350	B1670350 603292721
24	1096	69.0	812	12	B6716776	B6716776 602678182
25	1096	69.0	812	13	B1601551	B1601551 603250760
26	1096	69.0	965	14	B6977266	B6977266 AGENCOURT
27	1093	68.8	797	12	B1603668	B1603668 603251982
28	1090.5	68.6	797	13	B6715366	B6715366 602675503
29	1090	68.6	846	13	B1159757	B1159757 602863509
30	1089	68.5	796	13	BM042153	BM042153 603615790
31	1084	68.2	757	12	B6707750	B6707750 602671277
32	1084	68.2	794	13	B1601279	B1601279 603245010
33	1084	68.2	954	12	BF967543	BF967543 602287404
34	1082	68.1	790	13	B1551066	B1551066 603196425
35	1081	68.0	790	13	B1551811	B1551811 603192716
36	1077.5	67.8	891	13	B6708414	B6708414 602672374
37	1076.5	67.7	891	13	B1549292	B1549292 603189930
38	1070.5	67.4	802	13	B1458355	B1458355 603198705
39	1069	67.3	1100	14	BM914382	BM914382 AGENCOURT
40	1068.5	67.2	914	13	B1603523	B1603523 603244424
41	1068	67.2	748	13	B1553085	B1553085 603196058
42	1067.5	67.2	930	12	BF967857	BF967857 6028287176
43	1063.5	66.9	812	12	B6769968	B6769968 602739654
44	1063	66.9	688	13	B1544886	B1544886 603242328
45	1062.5	66.9	803	12	B6709360	B6709360 602673385

ALIGNMENTS

RESULT 1
LOCUS B1670367 1027 bp mRNA EST 12-SEP-2001
DEFINITION 603292738F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shluraki
 Toshynki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11790 row: 1 column: 09
 High quality sequence stop: 845.
 Location/Qualifiers

FEATURES source

1..1027
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5312024"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVA-3',
 size selected for average insert size 2.3 kb and
 normalized to R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 194 a 308 c 400 g 125 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.09e-124 Length: 1027
 Score: 1275.50 Matches: 290
 Percent Similarity: 90.248 Conservative: 6
 Best Local Similarity: 88.418 Mismatches: 21
 Query Match: 80.278 Indels: 12
 DB: 13 Gaps: 2

US-09-827-854-17 (1-317) x B1670367 (1-1027)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 |||||
 DB 26 ATGAAGTTCTGTGGCTGCTGCTGCTGCATCTCTGCGAGATGCCAGGCCAAGTGT 85
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluProGlnIns 40
 |||||
 DB 86 GACCAAGCGGTGGAGACAGCCGAGCCGAGCTGGCCAGACAGACCGAGTGGCAGAGC 145
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 |||||
 DB 146 GGGCAGCGGTGGAGACAGCCGAGCCGAGCTGGCCAGATCTGCGTGGTGGCAGACA 205
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 DB 206 CTGCTGAGAGAGGTGCAGGAGAGCTGCTCAGCTCCAGGTCAACCAGGAACAGAGGCGC 265
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaThrLysSerGluLeuGlnGluGlnIns 100
 |||||
 DB 266 CTGATGAGCAGACCATGAAGAGGCTTACAAATCGGAAGCTGGAGGAGCAACTG 325
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 DB 326 ACCCGGTTGGCGAGAGACGCGGCGAGCGCTCTCAAGAGCTGGAGCGGCGAGCGC 385
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
 |||||
 DB 386 CGGCTGGGCGGACATGAGAGCGTGTGCGCGCTGGTGCATACCGCGGCGAGGTG 445
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||

DB 446 CAGGCCATGCTGGCCAGACACCGAGAGCTGGGGTCCGCTCCGCTCCACCTGCC 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysCysLeuAlaValTyr 180
 |||||
 DB 506 AAGCTGGTGAAGGGGCTCCCTCCGGAGTCCGAGTACCTGACAGAAAGCGCTGGCAGTAC 565
 QY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-Gl 200
 |||||
 DB 566 CAGCGCGGGCCCGCGAGGGCGCGAGCGCGCTCAGCCCATCCGCGGCGCCCTGGGG 625
 QY 200 YProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPr 220
 |||||
 DB 626 CCCCCTGTGTGAACAGGTCCGCTGCGGGCCCGCCACTGTGTGGCTCCCTGGCCGCCACGC 665
 QY 220 OleuGlnGluArgAlaGlnAlaTrpGlyGluArg-LeuArgAlaArgMetGluGlu-Met 239
 |||||
 DB 686 GCTACAGAGACGGGGCCAAAGCCTGGGGCGAGCGGCTTGGCGCGCGGATGAGAGAGAG 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArgA 259
 |||||
 DB 746 GGCAGCGCGGAGCCCGGACCGCTGACCGAGTGAAAGAGACAGTGGCGGAGGTGCCGCG 805
 QY 259 IalysLeu---GluGluGlnAlaGlnGln-IleArgLeuGln-AlaGluAlaPheGlnAl 277
 |||||
 DB 806 CCCAAGCTGGACGAGCCAGCCAGCCAGCAGATACGCTGAGGCGCCGAGGCTTCTAGGC 865
 QY 277 AatGluLysSerTrpPheGluPro-----LeuValGluAspMetGlnArgGlnTrpAl 295
 |||||
 DB 866 CGGCGCTCAGAGAGTGGTGGTGAACCCCTGGGTGTGAACGACATGAGGCGCGGCGCAAGTGG 925
 QY 295 agilyLeuValGluLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSerA 315
 |||||
 DB 926 CGGCTGTGGAAAAAGGTGCAGGCTGCGGGGGCACACCGCGCTGTGTGCAGC-G 984
 QY 315 spasnHis 317
 |||||
 DB 985 ACAATCAT 992

RESULT 2

BG763371

LOCUS 602735433F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
 DEFINITION mRNA sequence.

ACCESSION BG763371.1 GI:14074024
 VERSION BG763371.1 GI:14074024
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 842)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M1718 row: b column: 10
 High quality sequence stop: 817.
 Location/Qualifiers

FEATURES source

1..842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4860585"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 154 a 250 c 337 g 101 t

ALIGNMENT SCORES:

Pred. No.: 6.01e-121
Score: 1239.00
Percent Similarity: 96.30%
Best Local Similarity: 95.56%
Query Match: 77.97%
DB: 12

Length: 842
Matches: 258
Conservative: 2
Mismatch: 9
Indels: 2
Gaps: 0

US-09-827-854-17 (1-317) x BG763371 (1-842)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 36 ATGAAGGTTCTGTGGCTCGTGTGTCACATTCTTGGCAGATGCCAGCCAAAGTG 95
QY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnThrGlnThrGlnThr 40
DB 96 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGACAGACGAGTGGCAGACC 155
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 156 GGCCAGCGCTGGGAATGTCACATGGTCTTTGGGATTAAGTGGGAGGAGGAGGAG 215
QY 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 216 CTGTCTGACAGAGTGGAGAGAGTGTCTCAGCTCCAGTCCAGCCAGGAACTGAGAGCGC 275
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaIleTyrLysSerGlnGlnGlnGlnGln 100
DB 276 CTGATGAGACAGACATGAGAGTGTGAAAGCTTCAAAATCGGAAGTGGAGAACAACTG 335
QY 101 ThrProValAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120
DB 336 ACCCGGCTGGGAGAGAGAGCGGCGACGCTGTCCAAAGAGCTGAGCGCGCGAGCGC 395
QY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 396 CGGCTGGGCGGAGACATGAGAGAGTGTGCGGCGCTGTGTCAGTACCGCGGAGAGTG 455
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
DB 456 CAGGCCATCTGGCGAGAGACCGGAGGCTGGCGGCTGCCCTCCCACTGCGCGC 515
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 516 AAGCTGCGTAAGCGGCTCTCGCGCATGACCTGACAAAGCGCTGCGCAGAGTAC 575
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 576 CAGGCGCGGCGCGGAGGCGCGCAGCGCGCTCAGCGCATCCGAGCGCGCGCGG 635
QY 201 ProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
DB 636 CCTTG-GTGGAACAGGCGCGGCGCGCGCAACTGTGGCTCCCTGCGCGCGCAGCG 694
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMet-61 240
DB 695 CTACAGGAGCGGCG 754
QY 240 yserArgTrpArgAspArgLeuAspGlnValLysGlnGlnGlnGlnGlnGlnGlnGln 260
DB 755 CAGCGGAGCG 814

QY 260 streuGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 269
DB 815 CTGGGAGGAGCAGCG 842

RESULT 3

LOCUS B1668318

DEFINITION B1668318 933 bp mRNA linear EST 12-SEP-2001
603295681p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5',
mRNA sequence.

ACCESSION B1668318
VERSION B1668318.1 GI:15582551
KEYWORDS EST.

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 933)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11798 row: a column: 21
High quality sequence stop: 780.

FEATURES

source

1..933
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5314844"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to R07 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 284 c 359 g 119 t

ALIGNMENT SCORES:

Pred. No.: 3.01e-120
Score: 1233.00
Percent Similarity: 93.38%
Best Local Similarity: 90.94%
Query Match: 77.60%
DB: 13

Length: 933
Matches: 261
Conservative: 7
Mismatch: 16
Indels: 4
Gaps: 1

US-09-827-854-17 (1-317) x B1668318 (1-933)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 75 ATGAAGGTTCTGTGGCTCGTGTGTCACATTCTTGGCAGATGCCAGCCAAAGTG 134
QY 21 GlnGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnThrGlnThrGlnThr 40
DB 135 GAGCAAGCGGTGGAGACAGAGCCGAGCTGGCCAGACAGACGAGTGGCAGACC 194
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60

Db 195 GGCACGCGTGGGACATGGCAGCTGGTCCCTTTGGATTACCTCCGCTGGGTGCAGACA 254
 QY LeuSerGluGluValGluGluGluLeuLeuSerSerGluValThrGluGluLeuArgAla 80
 Db 255 CTGTCTGACGAGGTGCAGAGAGACTGCTCAGCTCCCAAGTCCACCGAAGAACTAGGGCG 314
 QY LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
 Db 315 CTGATGACGACGACCATGAAAGAGTGAAGCTTCAAAATCGAAGTGAAGAGAACAACTG 374
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaAlaGluAla 120
 Db 375 ACCCGGGTGGCGAGAGACCGGGGCACTGCTCCAAAGAGCTGCAGGGCGCCAGAGCC 434
 QY ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 435 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCCCTGTGTGCACTACCGGGGAGAGTG 494
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 495 CAGGCGCATGCTCGCGCAGAGCACCAGAGAGCTCGGGTGGCCCTCCGCTCCACCTGCGC 554
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 555 AAGCTGGGTAAAGCGGCTCTCCGCGATGCCATCTGCAGAAAGCCCTGGCAGTGTAC 614
 QY 181 GlnAlaGluAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgGly 200
 Db 615 CAGCGCCGGGGCGCGAGAGGGCGCGAGCGGCTCAGCGCCATCCGAGAGCGCTGGGG 674
 QY 201 ProLeuValGluGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 675 CCCCTGTGTGAACAGGGCCGCGGCTCGCGCCCACTGTGGGCTCCCTGCGCGGACAGCG 734
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArg-MetGlu--GluMe 239
 Db 735 CTACAGAGAGCGGGCCAGGCTGTGGGAGAGGCTGGCGCGGATGGAGCAGCAATTG 794
 QY 239 TgLySerArgThrArgAspArgLeuAsp-GluValLysGluGlnValAlaGluValArg 259
 Db 795 GGGCAGCGCGGACCGCGGACCGCTGGAACGAGTGAAGAGAGAGCTGGGAAAGTGGCGC 854
 QY 259 IalysLeuGluGluGluGlnGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgL 279
 Db 855 GCCAAGGCTGGGGGAGCAGGCGCCGCAATTCGCTTGACAGGCCAA-GCCCTTCAGAGCGGCT 913
 QY 279 eulysSerTrpPheGlu 284
 Db 914 CCAAGAGCTGGTTCAA 930
 RESULT 4
 Bg761746 938 bp mRNA linear EST 15-MAY-2001
 LOCUS 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:481411 5'
 DEFINITION mRNA sequence.
 ACCESSION Bg761746
 VERSION Bg761746.1 GI:14072399
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (Bases 1 to 938)
 AUTHORS NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incycle Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LILM674 row: c column: 12
 High quality sequence stop: 767.
 Location/Qualifiers
 FEATURES
 source
 1. 938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:481411"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using Zap-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

BASE COUNT 171 a 272 c 374 g 121 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.8e-117 Length: 938
 Score: 1205.00 Matches: 270
 Percent Similarity: 91.95% Conservative: 4
 Best Local Similarity: 90.60% Mismatches: 13
 Query Match: 75.83% Indels: 11
 DB: 12 Gaps: 2

US-09-827-854-17 (1-317) x Bg761746 (1-938)
 QY 1 MetLysValLeuThrPpaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 57 ATGAAGGTTCTGTGGGTGGTGTGTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 116
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 117 GAGCAAGCGGTGGAGAGAGAGAGCGGAGCCGAGCTGGCCAGAGAGACCGAGTGGCAGAGC 176
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 177 GGCAGGCTGGGAAAGCGACATGGGCTCTTTGGATTACCTGGCTGGGGGAGACA 236
 QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 237 CTGTCTGACGAGGTGCAGAGAGAGCTGTCTCAGCTCCAGGTCAACCGAAGTGAAGGGCG 296
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlu 100
 Db 297 CTGATGACGACGACCATGAAAGAGTGAAGGCTTCAAAATCGAAGTGAAGAACTG 356
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 357 ACCCGGTGGCGAGAGAGAGCGGGCAGCGCTGTCCAAAGAGTGAAGAGCGGGCGAGGCC 416
 QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 417 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCCCTGTGTGAGTACCGCGGAGAGTG 476
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 477 CAGGCGCATGCTCGCGCAGAGACCGAGAGACTCGGGTGGCGCTCCCTCCACCTGCGCG 536
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 537 AAGCTGGGTAAAGCGGCTCTCCGCGATGCCGATGACCTGAGAAAGCGCTGCGCAGGTAC 596
 QY 181 GlnAlaGluAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgGly 200
 Db 597 CAGCGCGGGCGCGCAGAGGCGCGAGCGGCTGTCTCAGGCGCCATTCGCGAGCGCTGGG 656

QY 201 ProLeuValGluGlnGlyArgValAlaAlaThrValGlySer-LeuAlaGlyGlnPr 220
|||||
Db 657 CCCCTGTGGACAGAGGGCGCGTGGCGGCCGACACTGTGGCTCCCTTGCGCGGACGCC 716
QY 220 OleuGlnGluArgAlaGlnAlaTrp-GlyGluArg-LeuAlaAlaArgMetGlu---Glu 238
|||||
Db 717 GCTACAGGAGCGGGCCAGCGCTGGGGGAGCGGCTTGGCGCGGATGAGAGGACAT 776
QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValLys-GluGlnAlaAla---Glu 256
|||||
Db 777 TGGGACAGCCCGACCCCGCGACCGCTTGGACGAGGTGAAGGAGCAGGTGGCGGAGG 836
QY 257 ValArgAlaLysLeuGlnGlu-GlnAlaGlnGlnIleArgLeuGlnAla--GluAlaPhe 275
|||||
Db 837 TTGGGGCGCAAGCTGAGAGACAGGCGCCAGAGATACGCCCTGCAGCGCGGAGAGGCTTC 896
QY 276 GlnAlaArgLeuLysSerTrpPheGluProLeuVal 287
|||||
Db 897 CCAGCGCGGCTCAAAAGCTGTTCGAAAGCCCTGT 932
RESULT 5
B1600906 942 bp mRNA linear EST 07-SEP-2001
LOCUS B1600906 603249241P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
DEFINITION mRNA sequence.
ACCESSION B1600906
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 942)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@b-rrmail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LHAM11762 row: a column: 11
High quality sequence stop: 762.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5301010"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 177 a 284 c 361 g 120 t
ORIGIN
Alignment Scores: 5.18e-117 Length: 942
Pred. No.: 1202.50 Matches: 6
Score: 92.78% Conservative: 6
Percent Similarity:

Best Local Similarity: 90.72% Mismatches: 15
Query Match: 75.68% Indels: 8
DB: 13 Gaps: 1
US-09-827-854-17 (1-317) x B1600906 (1-942)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 73 ATGAAGTTCTGTGGGCTCGTTGCTGTCATATTCCTGGCAGATGCCAGGCAAGGTG 132
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 133 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCACAGACCGAGTGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
Db 193 GGCACGCTGGGAACTGGCTGCTGGCTCTTTGGATTTACTCTGCTGGGTGCAGACA 252
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTCTGACACAGTGGAGAGAGAGCTGCTACCTCCAGTCCACAGAGAACTAGAGGCGC 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
|||||
Db 313 CTGATGGACGAGACCATGAGAGTGAAGCGCTCAAAATCGGAATCGAGAGAACAACTG 372
QY 101 ThrProValAlaGlnGluThrArgAlaArgLysSerLysGluGlnGlnAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGAGAGAGCGCGGCTGTCCAAAGAGCTGAGGCGGCGAGGCGC 432
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 433 CGGCTGGGGCGGAGACATGAGAGAGCTGCGGCGCGCTGTGTACAGTACCGGCGAGGTG 492
QY 141 GlnAlaMetLeuAspGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCGATGCTCGGCGAGAGACCGGAGCTGCGGCTGCGCTGCCACCTCGCGC 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTAACGGCTCTCCGCGATGCGATGACCTGCAAGAGCGCTTGGCAGTGTAC 612
QY 181 GlnAlaGlyAlaArgGlyGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGGCGGGGGCGCGGAGGGCGCGAGCGCGGCTCAGCGCATCCGAGGCGCGTGGG 672
QY 201 ProLeuValGluGlnGlyArgValAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGGACAGAGGCGCGCTGGCGGC-GCCACTGTGGCTCCCTTGGCGCCAGCGCG 731
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAla-ArgMetGlu-GluMetG 240
|||||
Db 732 CTACAGAGAGCGGCGCGAGGCTGTGGGCGAGCGGCTGCGCGGATGAGAGCAATTTG 791
QY 240 LysArgThrArgAspArgLeuAsp-GluValLys-GluGlnAlaAlaGluVal-ArgA 259
|||||
Db 792 GAGCGGAGCGCGCGAGGCTGAGACGAGGAGCAAGGAGTTCGGAATGTCGCGG 851
QY 255 LysLysLeuGlnGlnAla---GlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 278
|||||
Db 852 CCACATTTGAGGAGACAGCGCCAGAGCAATATACGCTTTCAGGCGGAGGCTTCAGAGGCC 911
QY 278 rgleuLysSerTrpPheGluPro 285
|||||
Db 912 GCTCAAAAC-TGCTTGAACCTT 933
RESULT 6
BM042094 800 bp mRNA linear EST 07-NOV-2001
LOCUS BM042094 603615713P1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
DEFINITION mRNA sequence.
ACCESSION BM042094
VERSION BM042094.1 GI:16771361

Alignment Scores:

Pred. No.: 1,97e-115 Length: 927
 Score: 1187.50 Matches: 273
 Percent Similarity: 91.67% Conservative: 2
 Best Local Similarity: 91.00% Mismatches: 15
 Query Match: 74.73% Indels: 12
 DB: 12 Gaps: 1

US-09-827-854-17 (1-317) x BG472299 (1-927)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB ATGAAGGTTCTGTGGGCTCGTGTGTCATCTCTGCGAGATGCCAGGCAAGGTG 106
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 107 GAGCAAGCGGTGGAGACACAGCCGAGAGCCGAGCTCGCAGACAGCCAGTGGCGAGAC 166
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 167 GGGCAGCGCTGGGAATGGCAGCTGCTTTGGGATTACCTGCGGTGGGTCAGACA 226
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 227 CTGTCTGACAGGTGACAGAGGAGGAGTGTCTACCTCCAGCTCACCCAGGAACTGAGAGGCG 286
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 287 CTGATGACAGACACCATGTAAGAGGTGAAAGGCTTCAAAATCGGAATGAGAGCAACATG 346
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 347 ACCCGGTGGCGGAGAGACGCGGCGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGCAGGC 406
QY 121 ArgLeuGlyAlaAspMetLysPValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 407 CGGTGGGCGCGACATGAGACGCTGTGCGCGCTGTGTGACAGTACCGCGCGAGGTG 466
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 467 CAGGGCATCTCGCGCAGAGCACCGAGAGCTGCGGTGCGCTCCGCCCACTCGCGC 526
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
DB 527 AAGCTCGTAAGGGGCTCTCCGCGCATGCGCATGACCTCAGAAAGCGCTTGGCAGGTAC 586
QY 181 Gln-AlaGlyAlaArgGluGlyAlaGluArgGly--LeuSerAlaIleArgGluArgLeu 199
DB 587 CAGGGCGGGGCGCGGAGGCGCGCGCCCTCAAGGCCCATCCGCGCGGCCCTT 646
QY 200 --GlyProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyG 219
DB 647 GGGGCGCGGTGGGTAACAGGCGCGGTGCGGCGCGCCACTGTGGCTCTCCGTGGCGGCC 706
QY 219 LeuProLeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGln 239
DB 707 AGCG-CTACAGAGACCGGCGCCAGGCTGGGCGCAACGCTG-CGCGCGCGAGTGGAGAGA 764
QY 239 eGlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValAr 258
DB 765 TGGGGAGGAGGAGCCGCGAGCTGTGACCGAGGTGAAGAGAGAGTGGCGGAGGTGG 824
QY 258 GAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAla---PheGlnAl 277
DB 825 CCGCAAGCTGTGGAGGAAGCAGCGCAGATACGGCTGCAGCGCGAGGGGCTTCCAGGGC 884
QY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGluAspMet 290
DB 885 CCGGCTCAAGAGAGCTGTGAGAGCCCTGGGTGGAAGACATG 926

```

RESULT 8
 BI597743
 LOCUS

BI597743 922 bp mRNA linear EST 07-SEP-2001

DEFINITION 603248609f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
 mRNA sequence.
 ACCESSION BI597743
 VERSION BI597743.1 GI:15490682
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 922)
 NIH-MGC <http://mhc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiluyki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM11760 row: b column: 04
 High quality sequence stop: 782.

FEATURES
 source location/Qualifiers

1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5300259"
 /clone_1ib="NIH_MGC_96"
 /tissue.type="hypothalamus"
 /lab_host="DH10B"

/note="Organ: Brain; Vector: Bluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R07 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 176 a 273 c 359 g 114 t

ORIGIN

Alignment Scores:

Pred. No.: 4.05e-115 Length: 922
 Score: 1184.50 Matches: 255
 Percent Similarity: 92.83% Conservative: 4
 Best Local Similarity: 91.40% Mismatches: 16
 Query Match: 74.54% Indels: 5
 DB: 13 Gaps: 1

US-09-827-854-17 (1-317) x BI597743 (1-922)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 73 ATGAAGTCTGTGGGCTCGTGTGTCATCTCTGCGAGATGCCAGGCAAGGTG 132
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 133 GAGCAAGCGGTGGAGACACAGCCGAGAGCCGAGCTCGCAGACAGCCAGTGGCGAGAC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 193 GGCAGCGCTGGGAATGGCAGCTGCTTTGGGATTACCTGCGGTGGGTCAGACA 252
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 253 CTGTCTGACAGGTGACAGAGAGAGTGTCTACCTCCAGGTACCCAGGAAGTGAAGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100

```

Db 313 CTGATGACGACACATGAAGAGTTGAAGCCTACAAATCGGAAGTGGAGAACAACTG 372
 QY 101 ThrProValAlaIgluInuThrArgAlaArgLeuSerLysGluLeuGlnAlaIgluAla 120
 Db 373 ACCCCGGTGGCGAGGAGACGGGGCAGCGCTGTCCCAAGAGAGTGGAGGGCGGAGGCC 432
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 433 CGGCTGGCGCCGACATGAGGAGCGTGTGCGCGCGCTGTGACGATACCCGCGGAGGTG 492
 QY 141 GlnAlaMetLeuAspGlnSerThrGluLeuValArgLeuValArgLeuValSerHisLeuArg 160
 Db 493 CAGGCCATGCTCGGACGACGACCGAGAGCTGGGGTGGCTCGCTCCACCTCGCGC 552
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 553 AACCTGGTAAAGGGGCTCTCCGCGATGGCGATGACCTGCAGAAAGCGCCCTGGCAGTGTAC 612
 QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 613 CAGCGCGGGCGCGCGAGGCGCGCGAGCGCGCTC-AGCCCATCCGCGAGCGCTGGGG 671
 QY 201 ProLeuVal-GluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGln-P 220
 Db 672 CCCCTGGTGGAAACAGGGCGGGTGGCGCGCTGCTGCTCCCTGGCGGCGAGCC 731
 QY 220 roLeuGlnGluArg-AlaGlnAlaTyrPglLysArgLeuArgAlaArgMetGluGluMet 239
 Db 732 CGGTACAGAGAGGGGGCGGCGCTGGGGGAGCGGATACGGCGCGATGAGAGAGAGG 791
 QY 240 GlySerArgThrArgAspArgLeuAspGluValLysGlnValAlaGluValArgAla 259
 Db 792 GGCAGCGCGGACCCGAGACCCCTGAGACGAGTGAAGAGCAGGTGGCGGGGCGGAGCCA 851
 QY 260 LysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGln 276
 Db 852 CAAGCTGAAGAAACAGCCCAAAATACGCTTA---GCCGAGAGCTTCAGG 899
 RESULT 9
 BG774871 817 bp mRNA linear EST 15-MAY-2001
 LOCUS 602649975F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',
 DEFINITION mRNA sequence.
 ACCESSION BG774871
 VERSION BG774871.1 GI:14045188
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 817)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1612 row: e column: 20
 High quality sequence stop: 813.
 Location/Qualifiers
 1..817
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4760827"
 /clone_lib="NIH_MGC_40"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pORF7; Site_1: xhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GCCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 148 a 242 c 324 g 103 t
 ORIGIN

Alignment Scores:

Pred. No.: 3,06e-114 Length: 817
 Score: 1175.50 Matches: 250
 Percent Similarity: 96.55% Conservative: 2
 Best Local Similarity: 95.79% Mismatches: 7
 Query Match: 73.98% Indels: 4
 DB: 12 Gaps: 0

US-09-827-854-17 (1-317) x BG774871 (1-817)

QY 1 MetLysValLeuThrPalaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 40 ATGAGAGTTCTGTGGCTGTGGCTGTGGTCTGTGACATTCCTGGCAGATGCCAGGCCAAGTG 99
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 100 GACCAAGCGGTGGAGACAGAGCCGAGCCGAGCTCCGCGACAGACAGCCAGTGGCAGAGC 159
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 160 GGCACCGCTGGGAACATGGACATGGGCTCTTTGGATTACCTGGCGGGTGGAGACA 219
 QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 220 CTGTCTGAGCAGGTGCAGAGAGAGAGCTGTGCTGAGCTCCAGCTACCCAGCAACTGAGGGCG 279
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 Db 280 CTGATGAGACGAGACATGAAGAGTTGAAGGCTTACAAATCGAATCGAAGAGCAACTG 339
 QY 101 ThrProValAlaIgluInuThrArgAlaArgLeuSerLysGluLeuGlnAlaIgluAla 120
 Db 340 ACCCGGTGGCGGAGGAGACGGCGGCGGCTGTCCAAAGAGCTGAGCGGGCGGAGGCC 399
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 400 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGTGTCAGTACCGCGGCGAGGTG 459
 QY 141 GlnAlaMetLeuAspGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 460 CAGGCCATGCTCGGCGCAGAGACCGGAGAGCTGCGGCTCCCTCCACACCTGGCCG 519
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
 Db 520 AAGCTCGTGAAGGCGCTCTCCGCGATGCGAGTCTCCAGAAAGCGCTGGGAGAGTAC 579
 QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 Db 580 CAGCGCGGGCGCGGAGGAGGCGCGGAGCGGCTGAGGCCATCCGCGAGCGGCGGGG 639
 QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
 Db 640 CCCCTGGTGAAGAGAGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 698
 QY 221 LeuGlnGluArgAlaGlnAla--TrpGlyGluArgLeuArgAlaArgMetGluGluMetG 240
 Db 699 CTACAGAGAGCGGGCCAGGCGCTTGGGGAGCGGATGGCGCGCGGATGAGAGATGG 758
 QY 240 LysSerArgThrArgAspArgLeuAspGluValLysGlnGluValAlaGluValArgAla 259
 Db 759 GCCACCGGAGCCG--GAACCGCTGGAGAGAGTGAAGAGCAGGTGGCGGAGGTGCGCGCC 815

QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
Db 253 CTGTCTGAGAGGTGACGAGGAGGCTGCTGCAGTCTCCAGGTCCACGAGAACGAGGGG 312
QY 81 LeuMetAspGluThrMetIysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
|||||
Db 313 CTGATGGAGAGACCATGTAAGAGGTGAAAGGCTTACAAATCGGAATCGAGAACCACTG 372
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCGGGTGGAGGAGGAGGCGGCGGCTGCTCCAGAGCTGCAGGCGGCGGCGGCGG 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGTGTGCGGCGGCTGTGTGAGTACCGCGGAGGTG 492
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCAATGCTCGGCGGAGACACCGAGGAGCTGCGGCTGCGCTCCACCTGCGG 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTCAGAAAGCGCTGGCAGGTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlnValArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
|||||
Db 613 CAGCGCGGGGCGCGGAGGCGCGGCGGCTGTCAGGCGCATCGGCAAGCGCTCGGG 672
QY 201 ProLeuValGlnGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGTGAACAGAGGCGCGGCGGCGGCGGCTGCTGCGGCTGCGGCGGCGGCGG 732
QY 221 LeuGlnGlnArgAlaGlnAlaThrProGlyGlnArgLeuArgAlaArgMetGlnGlnLeu 239
|||||
Db 733 CTACAGAGACGGGCGGCGGCTGCGGCGGCGGCTG-CGGCGCGGATGAGAGAGATG 788
RESULT 13
BI551475 919 bp mRNA linear EST 05-SEP-2001
BI551475
LOCUS 603194314P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION BI551475
VERSION BI551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 919)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC Clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
Location/Qualifiers
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 9,01e-110 Length: 919
Score: 1134.00 Matches: 245
Percent Similarity: 92.88% Conservative: 3
Best Local Similarity: 91.76% Mismatches: 18
Query Match: 71.37% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-17 (1-317) x BI551475 (1-919)
QY 1 MetLysValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 73 ATGAAGTTCTGTGGGCTCGTGGTGTGTCACATTCCTGGCAGGATGCCAGGCCAAGGTG 132
QY 21 GlnGlnAlaValGlnThrLupProGlnLupProGlnLupArgGlnThrLupArgLupSer 40
|||||
Db 133 GAGCAAGCGGTGGAGACACAGCGGAGCCGAGGCTGCGCAGCAGACGACGAGTGGCAGAC 192
QY 41 GlyLupArgTPrpGluLeuAlaLeuGlnLupArgPheTPrpAspTyrLeuArgTPrpValGlnThr 60
|||||
Db 193 GGCCACGCTGGGAATGCGACTGGTGGCTTTTGGGATTAAGTGGCTGGGTTGGCAGATA 252
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
Db 253 CTGTCTGAGAGGTGACGAGGAGGCTGCTGCAGTCTCCAGGTCCACGAGAACGAGGCGG 312
QY 81 LeuMetAspGluThrMetIysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
|||||
Db 313 CTGATGGAGAGACCATGTAAGAGGTGAAAGGCTTAAATCGGAATCGAGAACCAACTG 372
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGGAGAGACGCGGCGGCTGTCCAGAGACTGAGCGGCGGCGGCGGCGG 432
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGGCGGACATGAGAGACGTGTGCGGCGGCTGTGTGAGTACCGCGGAGGTG 492
QY 141 GlnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCAATGCTCGGCGGAGACACCGAGGAGCTGCGGCTGCGCTCCACCTGCGGCGG 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr 180
|||||
Db 553 AAGCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTCAGAAAGCGCTGGCAGGTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlnValArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
|||||
Db 613 CAGCGCGGGGCGGAGTGGCGGCGGCGGCTGTCAGGCGCATCGGCAAGCGGCTCGGG 672
QY 201 ProLeuValGlnGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 673 CCCCTGTGTGAACAGAGGCGGCGGCGGCGGCTGCTGCGGCTGCGGCGGCGGCGGCGG 732
QY 221 LeuGlnGlnArgAlaGlnAlaThrProGlyGlnArgLeuArgAlaArgMetGlnGlnLeu 240
|||||
Db 733 A--CAGAGACGGGCGGAGGCTGGGGCGGAGCGGTGCGGCGGATGAGAGAAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
|||||
Db 791 CAGCGGAGACCGGAGACG--CTGGAGCAAGGTGAAGACAGCTGGGCGGAGAGG--TGGCCCA 847

Oy	260	steugluglucinalagin	266
	: : : :		
Db	848	GCTGAGGACCCAGCCAGAG	866
RESULT	14		
LOCUS	BM042228		
DEFINITION	757 bp	mRNA	linear
	603616186p1 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:5557004 5' ,		
ACCESSION	BM042228		
VERSION	BM042228.1	GI:16771495	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 757)		
TITLE	NIH-MGC http://mgc.nci.nih.gov//.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

Plate: LLCM2021 row: k column: 21
High quality sequence stop: 757.

FEATURES	Location/Qualifiers
source	1. .757

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/clone_id="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by Oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCGAG(c). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

Alignment Scores:	
Pred. No.:	7.83e-109
Score:	1124.00
Percent Similarity:	98.31%
Best Local Similarity:	98.31%
Query Match:	70.74%
DB:	13
Length:	757
Matches:	233
Conservative:	0
Mismatches:	3
Indels:	2
Gaps:	0

US-09-827-854-17 (1-317) x BM042228 (1-757)

1 MetLysValLeuTrpAlaLeuVal,ThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 50 ATGAAGTTCTGTGGGCTGCCGTGCTGGTCACATTCCTGGCAGGATGCCAGGCCAAGGTC 109

21 GUGNALAVALGUTHRGUPROGLUPLUARGINGINTHRGLTRPGINSER 40

20 110 GAGCGAGCCGATGGAGACAGAGCCCGAGGACCCGAGC 100

[illegible]

Oy	61	LeuSerGIuGlnValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgIa	80
Db	230	CTGTGTGAGCAGAGCTCAGAGAGAGCTGCTCAAGTCCACAGTCCACCAGAACTGAGGGCG	289
Oy	81	LeuMetAspGIuThrMetLeuSGluLeuLysAlaTyrTlysSerGIuLeuGIuGIuLeu	100
Db	230	CTGATGGACGAGACCATAGAGAGTTGAAGGCGTCAAAATCGGAATCTGAGGAACACTG	349
Oy	101	ThrProValAlaGIuGIuThrArgIaArgLeuSerTysGIuLeuGlnAlaAlaGlnIa	120
Db	350	ACCCTGGTGGCGGAGAGAACCGCGGGCAGCGCTGTCTCCAAAGAGCTCAGCGCGGAGGCC	409
Oy	121	ArgLeuGlnIaIaAspMetGIuAspValCysGlyArgLeuValGIuTyrArgGlyGIuVal	140
Db	410	CGGCTGGCGCGGACATGAGAGACTGTGCGGGCCCTCGTGTCAGATACCGCGGCAAGTG	469
Oy	141	GlnAlaMetLeuAspGlnSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	470	CAGGCGATGCTCGGCGACAGCACCGAGAGACTGCGGGTGTGCTGTCTCCACACTGGCC	529
Oy	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTyr	180
Db	530	AAGCTGCTTAGCGGGCTCTCCGCGATCCCATGCACTCGCAAGAGCGCTCGCATGTATC	589
Oy	181	GlnAlaGlyAlaArgGIuGIuAlaGlnArgGlyLeuSerAlaIleArgGIuArgLeuGly	200
Db	590	CAGGCGGGGCGCCGAGAGGCGCCGAGCGGGCTCAAGCGCCATCCGGCAGCGCTGGGG	649
Oy	201	ProLeuValGIuGIuGlnIaArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPr	220
Db	650	CCCCGTGGTGAACAGGGCGCGGTGCGGGCGGCCACTGTGGGCTCTTCGCGCGGACGCC	709
Oy	220	OLEuGlnGIuArgAlaGlnIaIleTPrGlyGIuArgLeuArgIaIleArgMet	236
Db	710	GCTACAGGAGCGGGC-CAGGCTCTGGGGGAGCGGGCTGGCGCGCGGAATG	757

RESULT 15
P0706130

LOCUS	BG706129	907 bp	mRNA	linear	EST 07-MAY-2001
DEFINITION	6025659093F1 NIH MGC 96 Homo sapiens cDNA clone IMAGE:4792030 5'				

miRNA sequence.
 BG706129
 ACCESSION

VERSION 20100420.1 01:40:00+0000
KEYWORDS EST.

[illegible]

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (base) 1 to 1007

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health. Mammalian Gene Collection (MGC)

JOURNAL unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

Toshiyuki and Piero Carninci (RIKEN)

DNA Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1AM10669 row: 1 column: 23
High quality sequence stop: 832

source	location/qualifiers
FEATURES	1. .907

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

```
/clone_lib="NIH_MGC_96"
```

```
/lab_host="DH10B"
```

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROR 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others
ORIGIN

Alignment Scores:

Pred. No.:	1.03e-107	Length:	907
Score:	1114.50	Matches:	242
Percent Similarity:	93.85%	Conservative:	2
Best Local Similarity:	93.08%	Mismatches:	12
Query Match:	70.14%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-17 (1-317) x BG706129 (1-907)

```
OY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db      |||||||
OY      75 ATGAAGTTCTGTGGCTGCTGTTGCTGTCACATTCTGGCAGATGCCAGCCAAAGTG 134
Db      |||||||
OY      21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db      |||||||
OY      135 GAGCAAGCGGTGAGACAGACGCCGAGCCGAGCTGCCACAGACCCAGATGGCAGACC 194
Db      |||||||
OY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db      |||||||
OY      195 GCCAGCGCTGGGAAGTGGCTGCTGCTTTGGATTACCTGGCTGGGTGCAGACA 254
Db      |||||||
OY      61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db      |||||||
OY      255 CTGTGTGACAGGTGCAGAGAGAGCTGCTCAGCTCCAGTCACCCAGGAAGTGAAGGCG 314
Db      |||||||
OY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
Db      |||||||
OY      315 CTGATGGAGACAGACCATGAGAGAGTTGAAGGCTTCAATCGAATCGAGAGACACTG 374
Db      |||||||
OY      101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
Db      |||||||
OY      375 ACCCGGTGGCGAGAGACGGCGGCTGCTCCAAAGAGCTGCAGCGCGGCGAGGC 434
Db      |||||||
OY      120 aArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVa 140
Db      |||||||
OY      435 CCGGCTGGGCGCGACATGGAGACGTGTGCGCGCTGTGTGACGTACCGCGGAGGT 494
Db      |||||||
OY      140 LglnAlaMetLeuAspGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuAr 160
Db      |||||||
OY      495 GAGAGCGCATGCTCGGCGACAGACCCAGAGAGCTGCGGCTGCGCTCCACCTGCG 554
Db      |||||||
OY      160 glyLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysCysLeuAlaValTrp 180
Db      |||||||
OY      555 CAAGCTGCGTAAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGCGAGTGA 614
Db      |||||||
OY      180 rGlnAlaGlyAlaArgGlnGlnGlnAlaArgGlyLeuSerAlaLleArgGluArgLeu-G 200
Db      |||||||
OY      615 CCAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGGCATCCGCGAGCGCTGGG 674
Db      |||||||
OY      200 LyrProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnP 220
Db      |||||||
OY      675 GCCCTGTGTGAGACAGGCGCGGTGCGGCGCCACTGTGGCTCCCTGCGNCGGCACG 734
Db      |||||||
OY      220 roLeuGlnGlnLarGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetG 240
Db      |||||||
OY      735 CCGTACAGAGCGCGCGAGGCTG---GGGCGAGCGGTGCGCGCGGATGAGAGATGG 791
Db      |||||||
OY      240 LysArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArg 258
Db      |||||||
OY      792 GCAG--CGGACCGG---ACGCTGACGAGTG--AAGAGCAGTGGGGAAGTGTGCGC 842
Db      |||||||
```

Search completed: March 14, 2003, 20:14:10
Job time : 1296.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.351 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVLMALIVFLAGCQAKV.....VEKYQAVGTSAAPVSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/US09827854/rumat_11032003_101610_27486/app_query.fasta_1.3576
-DB=GenEmbl -QEXT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTEXT=pto -NORM=ext -HEAPSIZE=5001 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn_1.1.13008.6runat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MAP -LARGEQUERT -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: gb.da:*
16: em.fun:*
17: em.in:*
18: em.jn:*
19: em.jnu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrl:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	99.2	1110	6 E00359	E00359 cDNA coding
2	1577	99.2	1110	6 E00823	E00823 DNA sequence
3	1577	99.2	1147	6 AX302545	AX302545 Sequence
4	1577	99.2	1156	6 BD004278	BD004278 Apo E hum
5	1577	99.2	1156	6 BD004278	BD004278 Apo E hum
6	1577	99.2	1156	6 BD004278	BD004278 Apo E hum
7	1577	99.2	1156	6 BD004278	BD004278 Apo E hum
8	1569	98.7	1157	6 AX333278	AX333278 Sequence
9	1568	98.7	1157	6 AX409597	AX409597 Sequence
10	1568	98.7	1157	6 AX409597	AX409597 Sequence
11	1568	98.7	1157	6 AX409597	AX409597 Sequence
12	1568	98.7	1157	6 AX409597	AX409597 Sequence
13	1565	98.5	1156	6 BD004279	BD004279 Apo E hum
14	1443	90.8	1178	6 MFAP0E	MFAP0E Monkey mRNA
15	1388.5	87.4	5491	6 AX358722	AX358722 Sequence
16	1388.5	87.4	41907	6 AF050154	AF050154 Homo sapi
17	1388.5	87.4	41907	6 AF050154	AF050154 Homo sapi
18	1388.5	87.4	41907	6 AF050154	AF050154 Homo sapi
19	1371	86.3	5515	6 HUMAPOE4	HUMAPOE4 Human apoli
20	1342	84.5	5413	9 AF261280	AF261280 Pan trogl
21	1259	79.2	4762	9 BABAP0E	BABAP0E Baboon apoli
22	1258	78.7	1138	4 AC021988	AC021988 Homo sapi
23	1250	78.7	1138	4 AC021988	AC021988 Homo sapi
24	1162.5	73.2	965	6 AX384545	AX384545 Sequence
25	1162.5	73.2	1108	4 BRAP0LPE	BRAP0LPE B.taurus mr
26	1162.5	73.2	5617	6 AX384541	AX384541 Sequence
27	1162.5	73.2	6026	6 AX384539	AX384539 Sequence
28	1162	73.1	1060	4 BABAP0L	BABAP0L B.taurus apoli
29	1144.5	72.0	1154	4 BRAP0EMR	BRAP0EMR B.taurus mr
30	1143	71.9	718	9 AF200499	AF200499 Pan trogl
31	1141	71.8	1045	10 M0SAPOE	M0SAPOE Mouse apoli
32	1141	71.8	1104	10 BC028816	BC028816 Mus muscu
33	1139	71.7	718	9 AF200506	AF200506 Gorilla g
34	1139	71.7	718	9 AF200506	AF200506 Pongo pyg
35	1135	71.4	718	9 AF200506	AF200506 Hylobates
36	1135	71.4	1122	4 SSAP0E	SSAP0E S.scrofa mr
37	1130	71.1	1126	6 AR164342	AR164342 Sequence
38	1130	71.1	1126	6 AR205885	AR205885 Sequence
39	1086	68.3	951	10 S76779	S76779 rAPOE
40	1064	67.0	959	10 M0SAPOE	M0SAPOE Mus musculu
41	1048.5	66.0	1069	10 RATAP0E	RATAP0E rat apoli
42	1021	64.3	228698	2 AC127479	AC127479 Mus muscu
43	1021	64.3	237653	2 AC073760	AC073760 Mus muscu
44	995	62.6	4856	10 M0SAPE	M0SAPE Mus musculu
45	994	62.6	4267	6 AR164387	AR164387 Sequence

RESULT 1

ALIGNMENTS

E00359 1110 bp RNA linear PAT 29-SEP-1997
LOCUS E00359
DEFINITION CDNA coding human apolipoprotein E3.
ACCESSION E00359
VERSION E00359.1 GI:2168646
KEYWORDS JP 1985118189-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1110)
Teranishi.Y., Takamatsu.N., Matsui.Y., Kimura.M. and Ikeda.Y.
REFERENCE DNa FRAGMENT
AUTHORS Patent: JP 1985118189-A 1 25-JUN-1985;
JOURNAL MITSUBISHI CHEM IND LTD
COMMENT OS human
PN JP 1985118189-A/1
PD 25-JUN-1985
PF 29-NOV-1983 JP 1983224980
PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
KIMURA MASAKO,
PI IKEDA YASUKO
PC C12N15/00,C07H21/04//C12P21/00;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue=livr;
FH Key Location/Qualifiers
FT CDS 15..968
FT sig_peptide 15..68 /product='apolipoprotein E3 precursor' FT
FT mac_peptide 69..965 /product='apolipoprotein E3 signal peptide' FT
FT Location/Qualifiers
FEATURES
source 1..1110
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 7.33e-86 Length: 1110
Score: 1577.00 Matches: 316
Percent Similarity: 99.688 Conservative: 0
Best Local Similarity: 99.688 Mismatches: 1
Query Match: 99.248 Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x E00359 (1-1110)
QY 1 MetIyValIeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 15 ATGAAGGTTCTGGGGCTGGCTGGTCACATTCCTGGCAGAGTCCAGGCCAAGGTG 74
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnProGlnSer 40
DB 75 GACCAAGCGGTGAGAGAGAGAGCCGAGCTGCGCCAGCAGAGCCGAGTGGCAGAGC 134
QY 41 GlnGlnArgTrpGlnLeuAlaLeuGlnArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 135 GGGCAGCGCTGGGAGACTGGGACTGGGCGCTTTGGGATTAAGTGGCGTGGCAGACA 194
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAlaGln 80
DB 195 CTGCTGAGCAGAGGTGAGAGAGAGCTGCTCAGCTCCAGGTCCAGCAAGTCAAGGCGC 254
QY 81 LeuMetAspGlnThrMetIyValGlnLeuAlaIyValThrIySerGlnLeuGlnGlnGln 100
DB 255 CTGATGAGCAGAGCCATGAAGAGGTTGAAGGCTTACAAATCGAAGTGGAGGACAAC 314

QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerIyGlnLeuGlnAlaGlnAla 120
DB 315 ACCCGCGTGGCGGAGAGAGAGCGGCGACGGCTGTCCAAAGAGCTGCAGCGCGGAGGCC 374
QY 121 ArgLeuGlnIyAlaAspMetGlnAspValCysGlnArgLeuValGlnTrpArgGlnVal 140
DB 375 CGCGTGGCGCGGACATGAGAGAGAGCTGTGGCGGCTGTGGCATGACCCGCGAGAGTG 434
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuAlaArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCGCATGCTGGCGCCAGAGACACCGAGAGACTCGGGTGGCGCTGCCCTCCACCTGGCG 494
QY 161 IyLeuCysIyAlaArgLeuLeuAlaArgAspAlaAspAspLeuGlnIyAlaValIyTr 180
DB 495 AAGCTGCGTAAGCGGGCTCCCGCGATGCGATGACCTGCAGAAAGCGCTGCAGAGTAC 554
QY 181 GlnAlaGlnIyAlaArgGlnGlnIyAlaArgIyLeuSerAlaIleArgGlnIyAlaGln 200
DB 555 CAGGCGGGCGCGCGAGGCGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGCGCTGGGG 614
QY 201 ProLeuValGlnGlnGlnIyArgValAlaArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
DB 615 CCCCTGTGGTAACAGAGGCGCGCTGCGGCGCGCCACTGTGGCTCCCTGGCGCGCGAGCGC 674
QY 221 LeuGlnGlnIyAlaArgAlaAlaIyTrpGlnIyAlaArgLeuAlaArgMetGlnIyMetGly 240
DB 675 CTACAGAGAGCGGCGCGAGCGCTGGCGAGCGCGCTGCGCGCGCGATGAGAGATGGCG 734
QY 241 SerArgThrArgAspArgLeuAspGlnValIyGlnGlnGlnValAlaGlnValAlaArg 260
DB 735 AGCCGAGCCCGCAGCGCGCTGAGAGAGTGAAGAGCAGAGTGGCGAGCGCGCGCGCAG 794
QY 261 LeuGlnGlnIyAlaGlnGlnIleArgLeuGlnAlaGlnIyAlaPheGlnAlaArgLeuIy 280
DB 795 CTGAGAGAGAGAGCCCAAGCAGATACGCTGCGAGCGCGAGCGCTCCAGCGCGCGCTCAAG 854
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlnIyLeuValGln 300
DB 855 AGCTGTTGAGAGCCCTCGTGGGAAAGATGACAGCGCCAGAGGCGCGCGTGGAGAGAG 914
QY 301 ValGlnAlaValAlaGlnIyThrSerAlaAlaProValProSerAspHis 317
DB 915 GTGAGAGCTGCGTGGGACACAGCGCGCGCTGTGCCAGCAGCAATGAC 965
RESULT 2
E00823 1110 bp DNA linear PAT 29-SEP-1997
LOCUS E00823
DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.
ACCESSION E00823
VERSION E00823.1 GI:2169084
KEYWORDS JP 1986096997-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1110)
AUTHORS Teranishi.Y., Matsui.Y., Ikeda.Y. and Kimura.M.
TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
MITSUBISHI CHEM IND LTD
COMMENT OS Human (Homo sapiens)
PN JP 1986096997-A/1
PD 15-MAY-1986
PF 16-OCT-1984 JP 1984216987
PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUKO, KIMURA MASAKO
PC C12P21/00,A61K35/74,A61K37/04,C12N15/00,(C12P21/00,C12R1:19),
PC (C12N15/00,
PC (C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue=livr;

```

FH Key Location/Qualifiers
FH 3'UTR 1..14
FT sig_peptide 15..68
FT /product='human apolipoprotein E signal FT
FT CDS peptide 69..968
FT /product='human apolipoprotein E' FT 3'UTR
FEATURES
source 1..1110
Location/Qualifiers
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN
1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
15 ATGAAGTTTGTTGGCTGGCTGGTGTGTCACATTCTCGCAGAGATGCCAGGCCAAGTGG 74
21 GtuglnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40
75 GACCAACGGGTGGAGACAGAGACCCGAGACCTGGCGCCAGAGACAGCCAGTGGCAGAGC 124
41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPvalGlnThr 60
135 GGCACGCGCTGGAGAACGTCGGTGGCTTTGGATTACCTGGCGCTGGTGGACACA 194
61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
195 CTGTCTAGCAGGTGCAGAGAGAGCTGTCTCAGCTCCAGGTACCCAGAACTGAGGGCG 254
81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
255 CTGATGACAGACCATGAAGAGTGAAGCCTTACAAATCGGAACGTGAGAAACAACCTG 314
101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
315 ACCCCGGTGGGAGAGAGACGGCGCTGTCCAAAGAGCTGCAGGGCGCGCAGGCC 374
121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
375 CGGCTGGGCGCCGACATGAGAGAGCTGTGGCGCCCTGTGCAATGCCGGCGAGGTG 424
141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
435 CAGGCCATGCTGGCCAGACACCGAGAGCTGGGGTGGCTCGCTGCCCTCCACCTGGCG 494
161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
495 AACCTGGTGAAGGCTCCCGCCGATGGCGATGACCTGGAGAAAGGCCCTGGAGTGTAC 554
181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGlyLysLeuGly 200
555 CAGGCCGGGCGCCGAGGGCGCCGAGCGGGCTTCAAGCCCATTCGCGAGGCCCTGGGG 614
201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
615 CCCTGGTGAACAGGGCGCGTGGCGGCCGACATGTGGGCTCCCTGGCGGCCGACGCC 674
221 LeuGlnGlnArgAlaGlnAlaThrPglyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
675 CTACAGAGAGGCGCCAGGGCTGGGGCGAGCGCTGGCGCGCGGATGAGAGATGGCG 724

```

```

241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
735 AGCGGAGACCCGCGACCGCTGGAGAGAGTGAAGAGACAGTGGCGGAGTGGCGGCAAG 794
261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
795 CTGAGAGAGAGCGCCAGACAGATACCTGACAGGCCGAGGCTCTCCAGGCCCTCAAG 854
281 SerTrpPheGluProLeuValGluAspMetGlnArgGlyTrpAlaGlyLeuValGlyLys 300
855 AGCTGGTTCAGCCCTGGTGGAGACATGACAGCCGAGTGGCGCGCTGGTGGAGAG 914
301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
915 GTGCAGGCTGCCCTGGGCAACACAGCGCCGCCCTGTGCCAGCAATCAAC 965
RESULT 3
AX302545 1147 bp DNA linear PAT 30-NOV-2001
LOCUS Sequence 63 from Patent WO0175177.
DEFINITION AX302545
ACCESSION AX302545
VERSION AX302545.1 GI:17383082
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.
AUTHORS Tumor markers in ovarian cancer
TITLE Patent: WO 0175177-A 63 11-OCT-2001;
JOURNAL THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES
source 1..1147
Location/Qualifiers
BASE COUNT 210 a 365 c 425 g 147 t
ORIGIN
1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
46 ATGAAGTTTGTTGGCTGGCTGGTGTGTCACATTCTCGCAGAGATGCCAGGCCAAGTGG 105
21 GtuglnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGlnThrGlnSer 40
106 GACCAACGGGTGGAGAACAGAGCCGAGACCTGGCGCCAGAGACAGACAGAGCTGGCAGAGC 165
41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPvalGlnThr 60
166 GGCACGCGCTGGAGAACACTGGCTGGCTTTGGATTACCTGGCGCTGGTGGACACA 225
61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
226 CTGTCTAGCAGGTGCAGAGAGAGCTGTCTCAGCTCCAGGTACCCAGAACTGAGGGCG 285
81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
286 CTGATGACAGACCATGAAGAGTGAAGCCTTACAAATCGGAACGTGAGAAACAACCTG 345
101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
346 ACCCCGGTGGCGAGAGAGACGGCGCTGTCCAAAGAGCTGCAGGGCGCGCGCAGGCC 405

```

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 406 CGCGTGGGCGCGGACATGAGAGAGCTGTGGCGCGCTGGGTGCATACCGGGGAGAGGTG 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 466 CAGCCCATGTGCTCGGCGCAGAGACCGAGAGGTGGGGTGGCCCTCGCTCCACCTCGGC 525
QY 161 LysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 526 AAGCTGGCTAAGCGGCTCTCTCCGGATGCCATGACCTGCAGAAAGCCCTGGCAGTAC 585
QY 181 GlnAlaGlyAlaArgGlnGlnGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 586 CAGCGCGGGGCGCGGAGGGGCGGAGCGCGGCTCAGCGCATCCGCGAGCGCTGGG 645
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 646 CCCCTGTGGAAAGGGCGCGCTCGGCGCCCACTGTGGCTCTCTGGCGCGCAGCCG 705
QY 221 LeuGlnGluArgAlaGlnAlaIleTyrGlyLeuArgLeuArgAlaArgMetGlnLysGly 240
|||||
Db 706 CTACAGAGAGCGGCGCCAGGCTGGGGGAGCGGCTGGCGCGCGGATGAGAGATGGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnAlaGluValArgAlaLys 260
|||||
Db 766 AGCGGAGAGCGCGGAGCGCTGGGAGAGCGGCTGGCGCGCGGATGAGAGATGGGC 825
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 826 CTGGAGAGAGAGCGCGCAGAGATAGCGCTGGAGGCGAGGCTTCGAGCGCGCTCAAG 885
QY 281 SerTyrPheGluProLeuValGlnLysAspMetGlnArgGlnTyrPheGlnGlyLeuValGlyLys 300
|||||
Db 886 AGCTGTTCAGAGCCCGCTGGGTGGAAGACATGCAGCGCCAGTGGCGGCTGTGTGAAGAAG 945
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 946 GTGAGAGCTGCGCGTGGGACACAGCGCGCGCTGTGCCAGCGACATTCAC 996
RESULT 4
BD004278 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004278 Apo E humanized mammal.
DEFINITION BD004278
ACCESSION BD004278.1 GI:18632239
VERSION JP 2001017028-A/2.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
TITLE Apo E humanized mammal.
JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
C12N15/09//C07K14/775,
PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC
FEATURES
source FH key Location/Qualifiers
FT CDS (61)..(1011).
1..1156 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
ORIGIN
Alignment Scores:
Pred. No.: 7 65e-86 length: 1156
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x BD004278 (1-1156)
QY 1 MetLysValLeuTrrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 61 ATGAAAGTTCTGTGGGCTGGGTTGCTGTGCATCTCTGGCAGAGATCCAGGCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGlnTyrPheGlnSer 40
|||||
Db 121 GAGCAAGCGGTGGAG 180
QY 41 GlyGlnArgThrPheGlnAlaLeuGlyArgPheTrrAspTyrLeuArgTrrPheValGlnThr 60
|||||
Db 181 GGCAGAGCTGGGAG 240
QY 61 LeuSerGlnGlnAlaGlnGlnGlnLysSerSerGlnValThrGlnGlnLysArgAla 80
|||||
Db 241 CTGTCTGAG 300
QY 81 LeuMetLysPheGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
|||||
Db 301 CTGATGAG 360
QY 101 ThrProValAlaGlnGlnGlnThrThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
|||||
Db 361 ACCCGGTTGGCGAG 420
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 421 CGGCTGGCGCGGACATGAG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLysAlaArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGCAATGCTCGGCGCAGAGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 161 LysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 541 AAGCTGTTCAGAGCCCGCTGGGTGGAAGACATGCAGCGCCAGTGGCGGCTGTGTGAAGAAG 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 601 CAGCGCGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCATTCGCCGAGCGCCCTGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220
|||||
Db 661 CCCCTGTGGAAAGGGCGCGCTCGGCGCCCACTGTGGCTCTCTGGCGCGCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleTyrGlyLeuArgLeuArgAlaArgMetGlnLysGly 240
|||||
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGGAGAGCGGCTGGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnAlaGluValArgAlaLys 260
|||||
Db 781 AGCGGAGAGCGCGGAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnLysPheGlnAlaArgLeuLys 280
|||||
Db 841 CTGGAGAGAGAGCGCGCAGAGATAGCGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 281 SerTyrPheGluProLeuValGlnLysAspMetGlnArgGlnTrrPheValGlnGlyLys 300
|||||
Db 901 AGCTGTTCAGAGCCCGCTGGGTGGAAGACATGCAGCGCCAGTGGCGGCTGTGTGAAGAAG 960

OY 301 valGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGACAGCTCCGTGGGACACGAGCCGCCCTGTGCGCCAGCAATGAC 1011
 RESULT 5
 HUMAPOE3
 LOCUS HUMAPOE3 1156 bp mRNA linear PRI 24-NOV-2000
 DEFINITION Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.
 ACCESSION K00396
 VERSION K00396.1 GI:178850
 KEYWORDS apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 355 to 1156)
 AUTHORS Breslow, J.L., McPherson, J., Nussbaum, A.L., Williams, R.W., Lojquist-Kahl, F., Karathanasis, S.K. and Zannis, V.I.
 TITLE Identification and DNA sequence of a human apolipoprotein E cDNA clone
 JOURNAL J. Biol. Chem. 257 (24), 14639-14641 (1982)
 MEDLINE 83082756
 PUBMED 6897404
 REFERENCE 2 (bases 250 to 777)
 AUTHORS Wallis, S.C., Rogne, S., Gill, L., Markham, A., Edge, M., Woods, D., Williamson, R. and Humphries, S.
 TITLE The isolation of cDNA clones for human apolipoprotein E and the detection of apoE RNA in hepatic and extra-hepatic tissues
 JOURNAL EMBO J. 2 (12), 2369-2373 (1983)
 MEDLINE 84131952
 PUBMED 6199196
 REFERENCE 3 (bases 1 to 1156)
 AUTHORS Zannis, V.I., McPherson, J., Goldberger, G., Karathanasis, S.K. and Breslow, J.L.
 TITLE Synthesis, intracellular processing, and signal peptide of human apolipoprotein E
 JOURNAL J. Biol. Chem. 259 (9), 5495-5499 (1984)
 MEDLINE 84185684
 PUBMED 6325438
 REFERENCE 4 (bases 88 to 1156)
 AUTHORS McLean, J.W., Elshourbagy, N.A., Chang, D.J., Mahley, R.W. and Taylor, J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 REFERENCE 5 (bases 577 to 624)
 AUTHORS Gill, L.L., Peoples, O.P., Pearson, D.H., Robertson, F.W., Humphries, S.E., Cumming, A.M. and Hardman, N.
 TITLE Isolation and characterisation of a variant allele of the gene for human apolipoprotein E
 JOURNAL Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985)
 MEDLINE 85279526
 PUBMED 2992507
 REFERENCE 6 (sites)
 AUTHORS Pail, S.C. Jr., Newhouse, Y.M., Clarke, H.R., Weisgraber, K.H., McCarthy, B.J., Mahley, R.W. and Bersot, T.P.
 TITLE III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant
 JOURNAL J. Clin. Invest. 83 (4), 1095-1101 (1989)
 MEDLINE 89198059
 PUBMED 2539388
 COMMENT [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant. [5] epsilon-2 allele. [6] sites: mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4,

FEATURES

epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3], [5]. The sequence shown is 57% homologous with human apo A-I and 81% homologous with rat apo E. For the epsilon-4 sequence, see the separate entry.
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)] and [3].
 Apo E is located on chromosome 19 -- Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1], [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA.
 Location/Organisms

source

1..1156

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="19q13.2"

/tissue_type="liver and blood"

1..1156

/gene="APOE"

61..1014

/gene="APOE"

/codon_start=1

/product="preapolipoprotein E"

/protein_id="AAB59546.1"

/db_xref="GI:178851"

/db_xref="GDB:600-119-691"

/translation="MKVLMALVTFLLAGCAQKQVAVTEPEPELRLQTEMSQRM EALGRFWDYLRWVTLSEVOEELSSQVVOELRALMDETRELKRAKSELEQLTP VAETGRRLSKELQAQAARLGDMDVCGRLVQYRGVQAMGOSTELRYLASHLR KLKRLRLDADLDLQRLAVYQAGAREGASAIREFRGLPVEGGRVRAATVGSAG QPLOEQAQMGERRARMEKSGSRTRDLDEYKEQVAVRAKLEEQAOQIRLQANAFO ARKSWPEPVEDMQRQAGLVEKVAVGISAARVPSDNH"

61..114

/gene="APOE"

/note="apolipoprotein E"

115..1011

/gene="APOE"

/product="apolipoprotein E mature peptide"

267

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"

/citation=[2]

/replace="c"

279

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

/citation=[4]

/replace="a"

288

/gene="APOE"

/note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]"

/citation=[4]

/replace="a"

369

/gene="APOE"

/note="in variant"

/citation=[4]

/replace="a"

409

variation

variation

variation

variation

variation

variation

variation

```

/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="a"
variation
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="c"
538
/ gene="APOE"
/ note="in type III hyperlipoproteinemia"
/ replace="t"
568
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="c"
586
variation
/ gene="APOE"
/ note="in epsilon-2 allele"
/ citation=[3]
/ citation=[5]
/ replace="t"
702
variation
/ gene="APOE"
/ note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]"
/ citation=[2]
/ replace="g"
783
variation
/ gene="APOE"
/ note="in variant"
/ citation=[4]
/ replace="t"
858
variation
/ gene="APOE"
/ note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3] and [4] variant; a in [4] normal."
/ citation=[4]
/ replace="a"
BASE COUNT      208 a      368 c      432 g      148 t
ORIGIN
Alignment Scores:
Pred. No.:      7,65e-86      Length:      1156
Score:          1577.00      Matches:      316
Percent Similarity: 99.68%      Conservative: 0
Best Local Similarity: 99.68%      Mismatches: 1
Query Match:    99.24%      Indels:      0
DB:             9      Gaps:      0
US-09-827-854-18 (1-317) x HDMAPOE3 (1-1156)
QY      1 MetIysValIeuTrpAlaIalaIeuValThrPheIeuAlaGlyCysGlnAlaIysVal 20
DB      61 ATGAAAGTTCGTGGGTGGTGGTGTGTCATTCCTGCGAGAGATGCCAGCCAAAGTG 120
QY      21 GluGlnAlaValaGluTrpGluProGluProGluLeuArgGlnGlnTrpGluTrpGlnSer 40
DB      121 GAGCAAGCGGTGAGAGACAGAGCCGAGACCTGCGCCAGAGACCGAGTGGCAGAGC 180
QY      41 GlyGlnArgTrpGluIleuAlaIeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      181 GGCACAGCGGTGGAGACAGCTGGGTCGCTTTGGGATTAACCTGCGCTGGTGAGACAGA 240
QY      61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB      241 CTGCTGAGCAGAGGTGACAGGAGGAGCTGCTCAGCTCCAGGTCAACCCAGAGACTGAGGGCG 300
QY      81 LeuMetAspGluThrMetIysGluLeuValaIaTyrIysSerGluLeuGluGlnIleu 100
DB      301 CTGATGACAGACCAATGAAGAGTTGAAGGCTTACAAATCGAGACTGGAGAAACAATG 360

```

```

QY      101 ThrProValaIaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaIaGlnAla 120
DB      361 ACCCGGTGGCGGAGAGAGACGGCGCACGGCTGTCCAAAGAGGTGACAGCGCGGAGGCC 420
QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyValArgLeuValGlnTyrArgGlyVal 140
DB      421 CGGCTGGCGCGGACATGAGAGAGAGTGTGGCGGCTGTGGCATGACCCCGGAGAGTG 480
QY      141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg 160
DB      481 CAGGCCATGCTCGGCCAGACACCGAGAGACTGCGGGTGGCGCTCGCTCCACACTGCCG 540
QY      161 LysLeuCysIysArgLeuLeuArgAspAlaAspAspLeuGlnIysArgLeuAlaValTyr 180
DB      541 AAGCTGCGTAAGCGGCTCTCCCGATGCCAGATGACCTGCAGAAAGCCCTGGCAGTGTAC 600
QY      181 GlnAlaGlyAlaArgGluGluValaArgGlyLeuSerAlaIleArgGluValArgLeuGly 200
DB      601 CAGCGCGGGCCCGGAGGCGCGGAGCGGCGCTCAGCCCATCCGCGAGCGCTGGGG 660
QY      201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB      661 CCCGTGTGACACAGGCGCGGTGGCGCGCCACTGTGGGTCCCTGGCCGCGGACGCCG 720
QY      221 LeuGlnIuArgAlaGlnAlaIaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB      721 CTACAGAGAGCGGCGCCAGGCTGTGGGAGACGGCTGCGCGCGGATGAGAGATGGCG 780
QY      241 SerArgThrArgAspArgLeuAspGluValaIysGluGlnValaIaGluValaArgAlaIys 260
DB      781 AGCGGAGCCCGCGCGCGCTGAGAGAGGTAGAGAGAGTGGCGGAGGTGCCGCCAAG 840
QY      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB      841 CTGAGAGAGCAGGCGCCAGCAGATACGCTGACGAGCGAGGCTTCACAGCGCCGCTCAAG 900
QY      281 SerTrpPheGluProLeuValaGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
DB      901 AGCTGTTCCAGGCCCTCGGTGGAGAACATGACCGCCAGTGGCGCGGTGTGGAGAG 960
QY      301 ValGlnAlaIaValaGlyThrSerAlaIaIaProValProSerAspHis 317
DB      961 GTGACAGGCTGCGTGGGACACAGCGCGCCCTGTGTGCCAGCACAATCATC 1011

```

RESULT 6
 BC003557 1186 bp mRNA linear PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, apolipoprotein E, clone MGC:1571 IMAGE:3355712, mRNA.
 ACCESSION BC003557 GI:13097698
 VERSION BC003557.1
 KEYWORDS MGC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1186)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 Info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Matheson, Candice Mcleavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Speed, Jeff Stott, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Zarr.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

source

Location/Qualifiers

1..1186

/organism="Homo sapiens"

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone="MGC:1571 IMAGE:3355712"

/tissue_type="Eye, retinoblastoma"

/clone_id="NH_MGC_16"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

46..999

/codon_start=1

/product="apolipoprotein E"

/protein_id="AAH03557.1"

/db_xref="GI:13097699"

/translation="MKVLAALVTFLLGCAKAEQAVTEPEPELROQTEMSQRM

ELAGREWDYLRWVOTLSEVOEELLSQVQELALNDELMKELAKKSELEQDLP

VAEETFRALSKELQAAQARLQADMEQVGRVQYGEVQAMIGSTEELRYLASHLR

KLKRLRLADDLQKRLAVYQAGAREGERSAIREFRLGVEEGRRVATVGSILG

QPLQERAAQMERLRLARMEGSRTRDLDEYKQAVRRAKLEQAOQIRLQAEAFQ

ARLSEPEPLVEDMDQRMAGLVKRYQAAVGSAPVPSDNH"

BASE COUNT 248 a 366 c 425 g 147 t

ORIGIN

Alignment Scores:

Pred. No.: 7,86e-86 Length: 1186
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-18 (1-317) x BC003557 (1-1186)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIaVal 20
DB 46 ATGAAGGTTCTGTGGCTCGTTGCTGGACATTCTCGCAGATGCCAGCCAAAGTG 105
QY 21 GlnGlnAlaValGlnThrLupProGluLeuArgGlnGlnThrLupProGlnSer 40
DB 106 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGTCCGACAGCAGCCAGTGGCAGAC 165
QY 41 GlnGlnAlaValGlnThrLupProGluLeuArgGlnGlnThrLupProGlnSer 60
DB 166 GGCACACCGTGGAACTGCTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 226 CTGTCTGACAGTCAGAGAGGAGTCTCCAGTCCAGTCCAGCAGCAACTGAGAGGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaIaIaIaIaIaIaIaIaIaIaIa 100
DB 286 CTGATGAGAGACCACTGAGAGAGTGAAGGCTTACAAATCGAACTGAGAGAACACTG 345
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaIaIaIa 120
DB 346 ACCCGCGTGGGAGAGAGCGGCGCGCTGTCCAAAGAGCTGAGGCGGCGCAGGCGC 405
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnThrArgGlnGln 140

DB 406 CGCTGGCGCGACATGAGAGACCTGTGCGCCCGCTGTGCTACCTACCGCGCAGAGTG 465
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGCGCATCTCTGCGCAGAGACACCGAGAGCTGCGGCTGCGCTCCCACTGCGC 525
QY 161 LysLeuCysAlaArgLeuLeuArgAlaAspAspLeuGlnLysArgLeuAlaValThr 180
DB 526 AAGCTGCGTAAGCGGCTCTCCGCGCATGTCAGTGCAGAGCGCTGCGCAGTGC 585
QY 181 GlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 586 CAGCGCGGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 645
QY 201 ProLeuValGlnGlnGlnArgValArgAlaIaIaIaIaIaIaIaIaIaIaIaIa 220
DB 646 CCCCTGTGTGAACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 705
QY 221 LeuGlnGlnArgAlaGlnAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 240
DB 706 CTACAGAGAGCG 765
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaIaIaIaIaIa 260
DB 766 ACCCGGACCG 825
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 826 CTGAGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 885
QY 281 SerTrpPheGlnProLeuValGlnLysAspMetGlnArgGlnTrpAlaGlnLys 300
DB 886 ACTGCTGTGACCG 945
QY 301 ValGlnAlaAlaValGlnThrSerAlaIaIaIaIaIaIaIaIaIaIaIaIaIa 317
DB 946 GTGCAAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 996
RESULT 7
LOCUS E08423 1110 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA coding human apolipoprotein E3.
E08423
ACCESSION E08423.1 GI:2176540
VERSION JP 1994315392-A/1.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1110)
Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.
AUTHORS Morimoto, H. and Teranishi, Y.
TITLE METHOD FOR PRODUCING APOLIPOPROTEIN
JOURNAL Patent: JP 1994315392-A 1 15-NOV-1994;
MITSUBISHI KASEI CORP
COMMENT OS Homo sapiens (human)
PN JP 1994315392-A/1
PD 15-NOV-1994
PE 11-JUN-1985 JP 1994015433
PI MORIMOTO HIRONORI, TERANISHI YUTAKA
PC C12P21/02, C07K13/00, C12N5/10, C12N15/12, C12N15/18, (C12P21/02,
PC C12R1:91),
PC (C12N5/10, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH key
FH Location/Qualifiers
FT source 1..1110
FT sig_peptide 15..68
FT mat_peptide 69..968
FT /product="human apolipoprotein E3".
FEATURES
source 1..1110
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 198 a 355 c 414 g 143 t
ORIGIN

Alignment Scores:
Pred. No.: 1.91e-85 Length: 1110
Score: 1570.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.80% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-18 (1-317) x E08423 (1-1110)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGGTTCTGTGGCTGCTTCTGTCTGTCATCTCTGGCAGATGTCAGGCAAGTG 74
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
DB 75 GAGCAACGGGTGGAGACAGACCGGAGCCGAGCTGGCCAGACCGAGCTGCGCAGAGC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGGCAGCGCTGGAGACTGGCACTGGCTGCTTTGGGATTACCTGGCTGGGTGCAGACA 194
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTAGCAGCGTGCAGAGAGAGACTGCTCAGTCCAGGTCACCCAGAACCTAGAGGCG 254
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 255 CTGATGAGCAGACCATGAAGAGAGTGAAGGCTTACAAATCGAAGTGAAGAGAACACTG 314
QY 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 315 ACCCGGTGGGGAGAGACGGGGCGGCTGTCACAGAGCTGCAGCGGGCGGCGAGGCC 374
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 375 CGGTGGCGCGGACATGAGAGAGAGTGTGGCGCGCTGTGCGATGACCCGCGGAGAGTG 434
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGCCCATGCTCGGCGCAGACACCGAGAGAGCTGGGGTGGCGCTCGCTCCACACTCGCG 494
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 495 AAGCTGGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCCCTGGCGAGTGTAC 554
QY 181 GlnAlaGlnAlaArgGlnGluGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 555 CAGCGCGGGGCGCGGAGGCGCGGAGCGGCGCTTCAGCGCCATCCCGAGCGCTTGGG 614
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 615 CCCCTGGTGGAAACAGGCGCGCTGGCGCGCCACCTGTGGGCTCCCTGGCGGAGCGCG 674
QY 221 LeuGlnGlnArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 675 CTACAGAGAGCGGCGCGGCTGGCGGAGCGGCTGGCGCGCGGAGTGAAGATGGGC 734
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 AGCGGAGACCGCGGACCGCTCGAGAGAGTGAAGAGAGAGTGGCGCGAGGTGGCGGCAAG 794
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 795 CTGAGAGAGAGAGCGCGCAGATACGCTTCAGAGCGAGCGCTTCAGCGCGCTCAAG 854
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300

DB 855 AGCTGTTCCAGCCCTGGTGGAGACATGACGCCAGTGGCGCGGCTGGTGGAGAG 914
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTGCAGGCTGGCGTGGCGACAGCGCGCGCTGTGCCAGCGACATTCAC 965

RESULT 8
BD004277 1156 bp DNA linear PAT 31-JAN-2002
DEFINITION Apo E humanized mammal.
ACCESSION BD004277.1 GI:18632238
VERSION BD004277.1 GI:18632238
KEYWORDS JP 2001017028-A/1.
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 1 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/1
PD 23-JAN-2001
PR 28-Apr-2000 JP 2000128919

PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10,PC
C12N15/09//C07K14/775
PC (C12N5/10,C12N1:91),C12N5/00,C12N15/00,(C12N5/00,C12N1:91) CC

FEATURES
source FH Key Location/Qualifiers
FT CDS (61)..(1011).
location/Qualifiers
1..1156
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 367 c 432 g 149 t

ORIGIN

Alignment Scores:
Pred. No.: 2.29e-85 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.74% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-18 (1-317) x BD004277 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGCTGCTTCTGTCTGTCATCTCTGGCAGATGTCAGGCAAGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
DB 121 GAGCAACGGGTGGAGACAGACCGGAGCCGAGCTGGCCAGACCGAGCTGCGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGAGACTGGCACTGGCTGCTTTGGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTAGCAGCGTGCAGAGAGAGTGTGCTCAGTCCAGGTCACCCAGAACCTAGAGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
DB 301 CTGATGAGACAGACCATGAAGAGTGAAGGCTTACAAATCGAAGCTGAGAGAACAACTG 360
QY 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120

Db 361 ACCCGGTGGCGAGAGACGCGGCGACGCGTGTCCAGAGAGCTGCAGCGCGCGAGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValIleThrArgGlyVal 140
Db 421 CGCGTGGCGCGACATGAGGAGCGGTGGCGCGCTGTGTGACATACCAGCGCGAGGTG 480
Qy 141 GAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTGGCCAGAGACACCGAGAGAGCTGGGGTGGCGCTCCGCTCCAGCTGGCC 540
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAAGCGCTCTCCCGCATGCCGATGACCTGCAGAGAGCTGGCAGAGTAC 600
Qy 181 GAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCGGGGCGCGAGGCGCGGAGCGGCTCAGCGCCATCCGCGCGCGCGCGGG 660
Qy 201 ProLeuValGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGTGTGAACAGGCGCCCGTGGCGCGCCACTGTGGCTCCCTGGCGCGCGAGCCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 721 CTACAGAGAGCGCGCCAGCGCTGGCGGAGCGGCTGGCGCGCGATGAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 781 AGCGGAGCCCGGACCGCTGGAGAGGTAAGAGACAGAGTGGGAGGCGCGCGCAAG 840
Qy 261 LeuGlnGluGlnAlaGlnAlaIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGCGGCCAGAGATACCGCTGTGCAGCGCGAGCCCTTCCAGCGCGCGCTCAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTTCAGCGCCCTGTGGTAAGACATGCACGCGCAGCGCGCGCTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspHis 317
Db 961 GTCCAGGCTGCGCGTGGCACACGCGCGCTGTGCCCGCAGCAATCATC 1011

RESULT 9
AX333278 1157 bp DNA linear PAT 09-JUN-2002
LOCUS Sequence 3787 from Patent WO0194629.
DEFINITION AX333278
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horigan, S., Soppet, D.R. and Weaver, Z.,
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 3787 13-DEC-2001;
JOURNAL Avalon Pharmaceuticals (US)
FEATURES
source 1. 1157
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 2,63e-85
Score: 1568.00
Percent Similarity: 99.058
Best Local Similarity: 99.058
Query Match: 98.688

Length: 1157
Matches: 314
Conservative: 0
Mismatches: 3
Indels: 0

DB: 6 Gaps: 0
US-09-827-854-18 (1-317) x AX333278 (1-1157)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGTTCTGTGGGCTGGGTGCTGTGTCACATTCCTGGCAGAGATGCCAGGCGAAGGTG 121
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnIleThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGAGACAGAGCGCGGAGCCCGACCTCGCGCAGAGACCGAGTGGCAGAGC 181
Qy 41 GlyGlnArgTrpLeuLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACGCGCTGGGAAGTGGCACTGGGTGGCTTTGGGATTAACCTGGCTGGGTGGAGACA 241
Qy 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTGAGAGAGGTGACAGGAGAGTGTACAGCTCCCAAGTCAACCAAGAACTGAGGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 302 CTGATGAGACGAGACCATGAAGGATTGAAGGCTTACAAATCGGAACTGAGAGAAACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTTAGCGAGAGAGACGCGGCGACGGCTGTCCAGAGAGCTGCAGAGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
Db 422 CGCGTGGCGCGGACATGAGGAGCGTGTGCGCGCGCTGTGTGACATACCAGCGCGAGGTG 481
Qy 141 GAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATCTCTGCGCAGAGCACCGAGGAGCTCGGGTGGCTCCCTCCCACTGGCGC 541
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTCGTAAAGCGGCTCTCCCGCATGATGACCTGCAGAGAGCGCTGGCAAGTGTAC 601
Qy 181 GAlaMetLeuGlyGlnSerThrGluLeuArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGCGCGCAGAGGCGCGCGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
Qy 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGTGTGAACAGGCGCGCGTGGCGCGCCACTGTGGCTCCCTGGCGCGCGAGCCG 721
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCCAGAGCTGTGGCGAGCGGCTCGCGCGGATGAGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 782 AGTCGAGACCGCGACCGCTGTGCAGGAGTGAAGAGACAGTGGCGGAGGTGGCGCGCAAG 841
Qy 261 LeuGlnGluGlnAlaGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGACAGGCCAGAGATACCGCTGCAGGCGCGAGGCTTCCAGGCGCGCTCAAG 901
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTTCAGAGCCCGTGTGTAAGACATGCACGCGCAATGAGCGCGCTGTGTGGAGAG 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspHis 317
Db 962 GTTCAGGCTGCGGTGGCACACAGCGCGCTGTGCCCGCAGCAATCATC 1012

RESULT 10
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 2244 from Patent WO0229103.
DEFINITION AX409597
ACCESSION AX409597

VERSION	AX009597.1	GI:21442302	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.		
TITLE	Gene expression profiles in liver cancer		
JOURNAL	Patent: WO 0229103-A 2244 11-Apr-2002;		
FEATURES	GENE LOGIC		
source	Location/Qualifiers		
	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="EMBL/GenBank Accession No. M12529"		
BASE COUNT	212 a	370 c	426 g
ORIGIN		149 t	
Alignment Scores:			
Pred. No.:	2.63e-85	Length:	1157
Score:	1568.00	Matches:	314
Percent Similarity:	99.05%	Conservative:	0
Best Local Similarity:	99.05%	Mismatches:	3
Query Match:	98.68%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-18 (1-317) x AX409597 (1-1157)			
OY	1	MettysValleuTPRPaAlaLeuDeuValThrpHeuAlaGlyCysGlnAlaLysVal	20
Db	62	ATGAAGTCTCTGTGGCGCTTCTGCTGCATTTCTCGGAGATGCCAGCCAAAGTG	121
OY	21	GlUglnAlaValaGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGGAGACAGACCAGGCCACAGCTGCAGCCAGCAGCAGAGTGGCAGAGC	181
OY	41	GlyGlnAArgTPRgLuLeuAlaLeuGlyAArgpHePrpAspTryLeuAArgTrpValGlnThr	60
Db	182	GGCCAGCGCTGGGAAGTGGCACTGGGTGGCTTTGGGATTAACCTGGCGTGGGTGACACA	241
OY	61	LeuSerGlnGlnValaGlnGlnGluLeuLeuSerSerGlnValaThrpGlnGluLeuAArgAla	80
Db	242	CTGTCTGAGCAGGTCGACAGAGAGAGCTGCTGATCCCAAGTCACCCAAAGACAGAGGCG	301
OY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu	100
Db	302	CTGAAGGACGAGACCATCAAGGATTTGAAGCCATCAAAATCGAAGTCGAGGAACAACGTG	361
OY	101	ThrpProValaGlnGluThrAArgAlaArgpLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362	ACCCCGGTAGGGAGAGAACCGGGGACAGGTGTCCAAGAGCTGCAGACGCGCCAGGCC	421
OY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyAArgLeuValGlnTrpArgGlyGluVal	140
Db	422	CGGGCGGGCGCGCATGTGAGAGAGTGTGGCGCCCTGTGGTCAGTACCGCGGAGAGTGTG	481
OY	141	GlnAlaMetLeuGlnGlnSerThrpGlnGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160
Db	482	CAGGCAATGCTGGCGCAAGACACCGAGAGAGCTGGGGGGCGCCCTGCCCTCCACCTCGC	541
OY	161	LysLeuCysLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTyr	180
Db	542	AAGTGGCTAAGCGGCTCTCTCGCGATCCCATATCACTGCAGAAAGCCCTGGCAGTGTAC	601
OY	181	GlnAlaGlyAlaAArgGluGlyAlaGluAArgLysLeuSerAlaLysLeuAArgGluAArgLeuGly	200
Db	602	CAGGCGCGGGCGCCGAGAGGCGCGAGACGCGGCTTCAGACGCATCCCGAGCGCTGGGG	661
OY	201	ProLeuValGlnGlnGlnArgValAArgAlaAlaThrpValGlySerLeuAlaGlyGlnPro	220
Db	662	CCCTGGTGGGAACAGGGCGGTGTCGGGCGCCCACTGTGGGTCTCCCTGGCGGCGCACCG	721

[illegible]

QY 121 ArgleuGlyAlaAspmetGluaspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
Db 422 CGCGTCGGCGGACATGAGGAGCGTGTGGCGCGCTGTGTGATGACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 482 CAGGCGCATCTCGCGCAGAGCACCGAGGAGTGGCGGTGGCTCCCTCCACCGTCGCGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 542 AAGCTCCGTAAGGGCTCTCCGCGATCCCGATGACCTCAGAAAGCGCTGCGCAGGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 602 CAGCGCGGGCGCGCGAGGGCGCGCGCGCTCAGCGCCATCCGCGACGCGCTCGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
|||||
Db 662 CCCCTGGTGGAAACAGGGCGCGCGCGCCACTGTGGCTCCCTGGCGCGCCAGCGG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyArgLeuArgAlaArgMetGluGluMetGly 240
|||||
Db 722 CTACAGGAGCGGGCCAGGCTGGGGCGAGCGCTGCGCGCGGATGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 782 AGTCGAGACCGGCGACCGCTGGAGCGAGTGAAGAGACAGTGGCGGAGGTGGCGCGCAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
|||||
Db 842 CTGGAGAGACAGGCGCGAGATACGCTCGCAGCGCGAGCCTTCAGCGCGCGCCCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 902 AGCTGTTGAGACCCCTGTGTGAGACATGACAGCGCACTGGGCGGCTGTGGAGAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 962 GGCAGAGCTGCGGTGGGACACAGCGCGCGCTGTGCCAGACACATATAC 1012
RESULT 12
HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
LOCUS Human apolipoprotein E mRNA, complete cds.
DEFINITION M12529
ACCESSION M12529.1 GI:178848
VERSION 1
KEYWORDS Homo sapiens (clone: pHAEl112,178,813). male 57-year old liver
SOURCE cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1157)
AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing
JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
MEDLINE 84212473
PUBMED 6327682
FEATURES
source location/Qualifiers
1..1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19q13.2"
/clone="pHAEl112,178,813".
/sex="male"
/tissue_type="liver"
/dev_stage="57-year old"
1..1157
/gene="APOE"
1..1157
mRNA

CDS
/gene="APOE"
/note="apoe mRNA; G00-119-691"
62..1015
/gene="APOE"
/note="precursor"
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAB59518.1"
/db_xref="GI:178849"
/db_xref="GDB:G00-119-691"
/translation="MKVLMALLVTFLLGCAKKYQAVTEPPEPRLROTEMSQRM
ELAGRFWDYLRWQTFLEQVQVEELSSQVQELRALMDETMKLAKVSELEQLTP
VAEETRARLSKELQTAQARLGADMEVCGRLVOYRGVQAMIGOSTEELRYLASHLR
KLKRLRLDPDDLQRLAVOAGAREGERSAIREIRLGLPVEGQRYRAATVGSILG
QIDERAQAMGERLRARAMEGSRTRDLQVKEQVAVRAKLEQAOQIRLQAEAFQ
ARLSWEPEPLVEDMQRQNAGLVEKVAIVGTSAPVPSDNI"
62..115
/gene="APOE"
/note="G00-119-691"
116..1012
/gene="APOE"
/product="apolipoprotein E"
/note="G00-119-691"
370
/gene="APOE"
/note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
410
/replace="g"
/gene="APOE"
/note="a in pHAEl112,1781; g in pHAEB13; G00-119-691"
569
/replace="g"
/gene="APOE"
/note="c in pHAEl112,1781; g in pHAEB13; G00-119-691"
784
/replace="g"
/gene="APOE"
/note="c in pHAEl112,1781; c in pHAEB13; G00-119-691"
859
/replace="c"
/gene="APOE"
/note="g in pHAEl112,1781; a in pHAEB13; G00-119-691"
variation
/gene="APOE"
/note="a in pHAEl112,1781; a in pHAEB13; G00-119-691"
variation
/gene="APOE"
/note="g in pHAEl112,1781; a in pHAEB13; G00-119-691"
variation
/gene="APOE"
/note="g in pHAEl112,1781; a in pHAEB13; G00-119-691"
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN 30 bp upstream of BstNI site.
Alignment Scores:
Pred. No.: 2.63e-85 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-18 (1-317) x HUMAPOE (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 62 ATGAAGTTCTTGTGGGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
QY 21 GlnGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
|||||
Db 122 GAGCAAGCGGTGGAGACAGAGCGGAGCCCGAGCTGCGCGACAGACCGAGTGGCAGAC 181
QY 41 GlyLnaArgTrpGluLeuAlaLeuGlnLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 182 GGCACAGCGCTGGAGACTGCGACTGCGCTTTGGAGTTACTCGCTGCGGTGGCAGACA 241
QY 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 242 CTGTGTGAGCAGGTGAGAGAGCTGCTCAGCTCCAGTACCCAGAACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100

|||||
Db 302 CTGATGACAGACCAAGAGGAGTGAAGCCCTACAAATCGGAACGGAGGAACAACCTG 361
Qy 101 ThProvalAlaGlugluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 120
Db 362 ACCCGGTACGAGGAGGAGAGCGGGCAGCGGTCTCCAGAGACTGCGAGAGCGCCAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGGCGCGACATGAGAGACGTGTGGCGCGCTGTGTGACGTACCGCGGAGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGCGCATGCTCGGCCAGACACCGAGAGACTGCGGGTGGCGCTCGCTCCACCTGCGC 541
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGGGTAAAGCGGCTCTCCCGATCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGCGCGCGAGGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCGCTGGGG 661
Qy 201 ProLeuValGluGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGAACAGGCGCGCGCGCGCGCCAGCTGTGGCTCCCTGCGCGCGCGCACCGC 721
Qy 221 LeuGluGluArgAlaGlnAlaTyrPglGlyArgLeuArgAlaArgMetGluGluMetGly 240
Db 722 CTACAGAGAGCGGCGCGCGCGCTGGGGGAGCGAGCGCTGCGCGCGCGATGGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 782 AGTCGGAGCCCGCAGCCCTCGAGAGGTGAAGAGAGAGAGTGGCGAGTGGCGCGCCAG 841
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGAGAGAGAGCGCCAGCAGATACGCTCAGAGCCAGGCGCTTCCAGGCGCGCGCTCAAG 901
Qy 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrPalaGlyLeuValGluLys 300
Db 902 AGTGGGTGAGCGCCCTGGTGGAGAACATGCACAGCGCGAGTGGCGCGCTGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGCTGGGCGACAGCGCGCGCTGTGGCCAGCGACATATCAC 1012
RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279 Apo E humanized mammal.
DEFINITION BD004279
ACCESSION BD004279.1 GI:18632240
VERSION JP 2001017028-A/3.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1156)
Apo E humanized mammal
Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
AO1k67/027,A61k45/00,A61p25/28,A61p43/00,C12N5/10, PC
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

.. Alignment Scores:
Pred. No.: 3,97e-85 Length: 1156
Score: 1565.00 Matches: 315
Percent Similarity: 99.378 Conservative: 0
Best Local Similarity: 99.378 Mismatches: 2
Query Match: 98.49% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x BD004279 (1-1156)
Qy 1 MetLysValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGACAGTCTGTGGGCTGCTGTGTGTCATCTCTGGCAGAGATGCCAGGCAAGGTG 120
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGAGCGGAGCCGACCTGCGCAGACAGACCGAGTGGCAGAGC 180
Qy 41 GlyLAlaThrPglLeuLeuAlaLeuGlyArgPheThrAspTyrLeuArgTyrPalaGlnThr 60
Db 181 GCGCAGCGCTGGAGACTGCGCAGCTGCTGCTTTGGGATTTACTGCGCTGGGTGCAGACA 240
Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGAGCAGAGTGCAGAGAGAGTGTCTGACCTCCAGATCCAGAGTAAGTGAAGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGCAGACCAATGAAGGAGTGAAGCCTTACAAATCGGAATGGAGGAACAACATG 360
Qy 101 ThProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGAGGAGAGAGCGCGGCGAGCGCTGTCCAGAGAGCTGCAGCGCGCGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGCGACATGAGAGACGTGGCGCGCGCTGTGTGACGTACCGCGGAGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGATGCTCGGCCAGAGACACCGAGAGACTGCGGGTGGCTCGCTCCACCTGCGC 540
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTGAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCGCTGCAGATGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlnGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCGCGAGGCGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGluGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 661 CCCCTGGTGAACAGGCGCGCGCGCGCGCGCTGTGGCTCCCTGGCGCGCGCGAGCGCG 720
Qy 221 LeuGluGluArgAlaGlnAlaTyrPglGlyArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGCGCGCGCGCTGGGGGAGCGCGCTGCGCGCGCGATGGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGACCGCGAGCGCTGTGAGAGAGTGAAGAGACAGTGGCGGAGCTGGCGCGCAAG 840
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||

Db 841 CTGGAGGACAGCCAGCATATCGCTGACAGCCGCTTCCAGCCCGCCTCAAG 900
Qy 281 serTrrpHegluProleuValgluaspMetgluArggluTrpAlaGlyLeuValgluLys 300
Db 901 AGCTGTTGACCCCTGTTGGTAGAGCATGCAGCGCATGTGGCCGCTGTTGGAGAA 960
Qy 301 ValgluAlaValgluIleThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGACAGGCTGCGTGGCAGCAGCGCCGCTGTGCTCCAGCAGCATATCAC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS Monkey mRNA for apolipoprotein E.
DEFINITION X13887
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti,K.R.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and
Melchior,G.W.
TITLE Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
Location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/tissue_type="liver"
/clone_idb="lambda gt11"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA 18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLMALVTFLLGCOAKVEQVPEPEPELROQAGSGQPM
ELALGRFWDYLRWQTLSEYOELISPOVQELTTLMDETMKELKAKSELEBOLSP
VAETRRARLSKELQAAOARLGADMEDVSRLOYRSEVOAMLGOSTEELRARLSHHR
QLKRLRLDADDLQRLAVQAGAREGAREVSAIRELGLVDEGRARRATVGSILAS
QPLDERAOLSERLRARKMEGSRTRDLDEKEDVALEVRKLEQAOQISLQNEAFQ
ARKLSWEFLVEDMOKWAGLVEKVOAAGVASTAVPIDNH"
sig_peptide
mat_peptide
misc_feature
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 7.49e-78
Score: 1443.00
Percent Similarity: 93.69%
Best Local Similarity: 93.06%
Query Match: 90.81%
DB: 9
US-09-827-854-18 (1-317) x MFAPOE (1-1178)
Qy 1 MetLysValLeuTrpAlaIleuLeuValIThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 83 ATGAAGGTTCTGTGGGCTCGTTGGCTGTCATCTCTGGCAGAGATGCCAGGCGCAAGGTG 142
Qy 21 GluGlnAlaValgluThrGluProgluProgluLeuArgGluGluThrGluTrpGlnSer 40
Db 143 GAGCAACCGGTGAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 202
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 203 GGCCAGCCCTGGGAGAGTGGACACTGGCTCTTTGGGATTTACTCGCTGGGTGGAGACA 262
Qy 61 LeuSerGluGluValgluGluLeuLeuLeuSerSerGluValThrGluGluLeuArgAla 80
Db 263 CTGTCTGACAGAGTGCAG 322
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlu 100
Db 323 CTGATGAGACGAGACCATGAAAGAGTTGAAGGCTTCAAAATCGGAAGTGAAGAGACGTG 382
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluGlnAlaAla 120
Db 383 ACCCGGTGGCGAG 442
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTrpArgGlyGluVal 140
Db 443 CGGTGGGTGCGGACATGAG 502
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGCATCTGGGCGAGAGTACCGGAGAGTGGGGGCGGCTCCCTCCACCTGGCGC 562
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGluLysArgLeuAlaValTyr 180
Db 563 AAGCTTCGCAAGCGGCTCTCCGCAATGCTGATGACCTCAGAGAGGCTTGGCAGTGTAT 622
Qy 181 GlnAlaGlyAlaArgGluGluValgluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 623 CAGGCGGGGCGCGAG 682
Qy 201 ProLeuValgluGluGluArgValArgAlaAlaThrValgluSerLeuAlaGluPro 220
Db 683 CCCCTGGGTGAGACAGAGGCGCGTGGCGGCGGCACACTGTGGCTCCCTGGCCAGCCAGCG 742
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 743 CTTCAGAGAGCGGCCAGGCTTGGGTGAGCGGCTTCCGCGACGATGAGAGAGATGGGC 802
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 803 ACCCGGAGCCGCGAGCGCGCTGGAGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 862
Qy 261 LeuGluGluGlnAlaGluGlnIleArgGluGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 863 CTGGAGGAACAGCGCCAGAGATAGCTCAGCGCGAGAGGCTTCCAGGCGCGCTCAAG 922
Qy 281 SerTrpPheGluProLeuValgluAspMetgluArggluTrpAlaGlyLeuValgluLys 300
Db 923 AGCTGTTGACCCCTGTTGGTAGAGCATATGCAGCGCATGTGGCTGGGTGGAGAA 982
Qy 301 ValgluAlaValgluIleThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGACAGGCTGCGTGGGCGCAGCAGCGCCCTGTGCTCCATGCAGCATATCAC 1033
RESULT 15
AF261279
LOCUS AF261279 5491 bp DNA linear PRI 27-OCT-2000
DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5491)
AUTHORS Nickerson,D.A., Taylor,S.L., Fullerton,S.M., Weiss,K.M.,
Clark,A.G., Stengard,J.H., Salama,V., Boerwinkle,F. and Sing,C.F.
TITLE Sequence diversity and large-scale typing of SNPs in the human
JOURNAL apolipoprotein E gene
MEDLINE Genome Res. 10 (10), 1532-1545 (2000)
PUBMED 20499366
11042151
REFERENCE 2 (bases 1 to 5491)
AUTHORS Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Department of Molecular Biotechnology,
University of Washington, Box 357730, Seattle, WA 98195, USA
FEATURES
source 1..5491
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
/contig=">108"
/note="putative"
repeat_region /rpt_family="MIR"
/rpt_type="dispersed"
73
/frequency="0.01"
/replace="t"
variation
repeat_region /note="putative"
/rpt_family="MIR"
/rpt_type="dispersed"
308
/frequency="0.01"
/replace="t"
satellite /note="putative"
/rpt_type="tandem"
/note="putative"
repeat_region /note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
471
/frequency="0.01"
/replace="g"
variation
repeat_region /frequency="0.01"
/replace="t"
560
/frequency="0.22"
/replace="t"
variation
satellite /frequency="0.07"
/replace="c"
/note="putative"
/rpt_type="tandem"
832
/frequency="0.45"
/replace="t"
variation
repeat_region /product="apolipoprotein-E"
1163
/frequency="0.35"
/replace="c"
1522
/frequency="0.01"
/replace="a"
1575
/frequency="0.01"
/replace="t"
variation
repeat_region /note="APOE"
/codon_start=1

/product="apolipoprotein-E"
/protein_id="AAG27089.1"
/db_xref="GI:11034801"
/translation="KKVLAALALVFLAGCCQAKVBOAVETEPEPELROTEWOSGQRM
EIALGFPMYLERWOTLSROVDELLISQVDELRLADETKKELKAKSELEQLT
VAEETRLRSKELQAAQRLCADMEDVCRLVQYGEVQAMIGOSTBELRVLAHLR
KLRRLRLRDADLQRLAVYQAGARGARGLSAIRELGPLVEGRKRAATVGSILAG
QPIQERAAQMGEMRLRAMEGMSRTDRIDEVKEQVAVERRAKLEEQAOQIRLQAEAFQ
ARLKSWEPEFLVEDMOROMAGLVEKVOAAVGTSAAPPSDNH"
1998
/frequency="0.10"
/replace="a"
/note="putative"
repeat_region /rpt_family="Alu"
/rpt_type="dispersed"
2440
/frequency="0.21"
/replace="a"
/note="putative"
variation /rpt_family="Alu"
/rpt_type="dispersed"
2907
/frequency="0.01"
/replace="g"
variation
repeat_region /frequency="0.01"
/replace="c"
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
3673
/frequency="0.01"
/replace="g"
variation
repeat_region /note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
3673
/frequency="0.01"
/replace="g"
variation
repeat_region /frequency="0.01"
/replace="cl"
3937
/frequency="0.14"
/replace="c"
variation
repeat_region /frequency="0.01"
/replace="t"
4075
/frequency="0.07"
/replace="t"
4735
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
5065
/note="putative"
/rpt_family="Alu"
/rpt_type="dispersed"
5229
/frequency="0.03"
/replace="gg"
variation
repeat_region /frequency="0.07"
/replace="t"
5229
/frequency="0.40"
/replace="g"
variation
repeat_region /frequency="0.13"
/replace="a"
5361
/frequency="0.06"
/replace="c"
variation
repeat_region /product="apolipoprotein-E"
1163
/frequency="0.35"
/replace="c"
1522
/frequency="0.01"
/replace="a"
1575
/frequency="0.01"
/replace="t"
variation
repeat_region /note="APOE"
/codon_start=1

Alignment Scores:

Pred. No.: 6,8e-74
 Score: 1388.50
 Percent Similarity: 60.60%
 Best Local Similarity: 60.60%
 Query Match: 87.38%
 DB: 9

Length: 5491
 Matches: 303
 Conserved: 0
 Mismatches: 3
 Indels: 14
 Gaps: 1

US-09-827-854-18 (1-317) x AF261279 (1-5491)

```

QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3003 TTCCACAGACAGATGCCAGGCGCAAGCGTGAGCAAGCGTGAGACAGAGCGGAGCCCGAG 3062

QY 32 LeuArgGlnGlnThrGluTrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPhe 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3063 CTGCGGCGAGAGACCGACAGGTGGCAGAGCGCCAGCGCTGGAGACTGGACCTGCGCTTT 3122

QY 52 TrpAspTyrLeuArgTrpValGlnThrLeuSerGlyGlnValGlnGluLeuLeuSer 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3123 TGGGATTACCTGGCTGGGTGGGCGACACTGTGACAGAGTGCAGAGAGAGTGTCTCAGC 3182

QY 72 SerGlnValThrGlnLeu----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3183 TCCAGGTACACCGAGAACTGAGTGTAGTGTCCCATCTGCGCCCTTGACCCCTCTGTGTC 3242

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3243 GCGGCTATACCTCCCGAGGTCCAGGTTCATTCTGCGCCCTGTGCTTAAGTCTTGGGGG 3302

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3303 CCGGGTCTCTGCTGTGCTTCTAGCTTCTCTCCATTCTGACTCCTGGCTTACGCTTC 3362

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3363 TGGAAATCTCTCTCAGCTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3422

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3423 CGTCTGGCTCTGTCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3482

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3483 CTCACGTGTGTGCCAGGCTGTGTGAATCTCTGGGCTCAAGCATCTCCGCGCTCGG 3542

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3543 CCTCCCAAGTGTGGGATTAGAGCATAGCACCTTGCGCGCTCTAGCTCTCTCT 3602

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3603 TCGTCTGTGCTCTGCTCTGTGATCTGTCTGTGATCTGTCTGTCTGTCTCTCTCTCT 3662

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3663 GCCTGCGCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3722

QY 78 ----- 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3723 CCCCATCCAGCCCTTCTCCCGCTCCCATGTGAGACACCTCCCGCCCTCTCGGCG 3782

QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 98
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3783 CAGGGCGCTGATGAGACGACATGAAAGGCTTACAAATCGGAATCGGAGTGAAGA 3842

QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3843 ACAATGACCCCGGTGGGAGAGACGCGGCGACGCTGTCCAGAGAGCTGACGCGCG 3902

QY 118 aglnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 3903 GCAGGCCCGGCTGGCGGACATGGAGAGAGTGTGCGGCCCGCTGTGCTACATCCGCG 3962

QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerH 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3963 CAGAGTGCAGAGCATCTCGGCCAGAGCACAGAGAGTGTGGGTGCGCTCGCTCCCA 4022

QY 158 sLeuArgLysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4023 CTTGCCAGAGCTCGTAAAGGGGCTCTCCGCGATGCGATGACCTGCAGAAAGCGCTGGC 4082

QY 178 aValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaLeuArgL 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4083 AGGTACCAAGCGCGGGCGCGAGGCGCGCGAGCGCGCGCTTCAGCCATCCGAGAGCG 4142

QY 198 gLeuGlyProLeuValGluGlnGlyArgValArgAlaLeuThrValGlySerLeuAla 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4143 CTTGGGGCCCTGTGTGAACAGGCGCGCTGTGCGCGCCACTGTGTGGCTCCTGTGCGG 4202

QY 218 yGlnProLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlu 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4203 CCAGCGGCTACAGAGACGCGGCCAGGCTGTGGCGAGCGGCTCGCGCGGATGAGAGA 4262

QY 238 uMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4263 GATGGCACCGCGGACCGCGCGCTGTGACGAGGTGAAGAGACAGGTGGCGAGGTGCG 4322

QY 258 gAlaLysLeuGluGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAla 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4323 CGCCAAAGCTGGAGAGGACGAGCCAGCATACCTGTGAGGCGGAGCGCTTCCAGGCGCG 4382

QY 278 gLeuLysSerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4383 CCTCAAGAGCTGTGTGAGCGCTGTGTGAAGCATGACAGCGCATGTGGCGCGGCTGGT 4442

QY 298 lGluLysValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4443 GGAGAGGTGACAGCTGTGGGACACAGCGCGCGCTGTGCTCCAGCGACATATCAC 4500

```

Search completed: March 14, 2003, 17:32:43
 Job time : 2314.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 : Search time 178.728 Seconds

(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVLMAALLTVFLAGCOAKV.....VEKVOAAGVTSAPVPSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPto_spool/US09827854/runat_11032003_101609_27476/app.query.fasta_1.3576
-DB=N.Geneseq_101002 -QFMT=fastap -SUFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854 @cgn2.1.1.1201 @runat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG1SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-VGAPOP=10 -VGAPEXT=0.5 -DELop=6 -DELExt=7

```

Database :

1:	N.Geneseq_101002.*
2:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1589	100.0	1156	24	AAD22051	Human apolipoprotein E
2	1577	99.2	954	24	AAD26035	Human apolipoprotein E
3	1577	99.2	1110	7	AA060409	Human apolipoprotein E
4	1577	99.2	1147	22	AB083113	Human apolipoprotein E
5	1577	99.2	1156	22	AA083115	Human apolipoprotein E
6	1577	99.2	1156	24	AAD22048	Human apolipoprotein E
7	1573	99.0	1156	24	AAD22052	Human apolipoprotein E
8	1569	98.7	1156	22	AA084314	Human apolipoprotein E
9	1569	98.7	1156	24	AAD22049	Human apolipoprotein E
10	1568	98.7	1157	17	AA060957	Human apolipoprotein E
11	1568	98.7	1157	24	AB083115	Human apolipoprotein E
12	1568	98.7	1157	24	AB083115	Human apolipoprotein E
13	1568	98.7	1157	24	AB083115	Human apolipoprotein E
14	1567	98.6	1110	6	AA050450	Human apolipoprotein E
15	1565	98.5	1156	22	AA084316	Human apolipoprotein E
16	1565	98.5	1156	24	AAD22047	Human apolipoprotein E
17	1562	98.3	1156	24	AAD22050	Human apolipoprotein E
18	1473.5	92.7	1279	22	AA022437	Human apolipoprotein E
19	1454.5	91.5	1107	19	AA075756	Human apolipoprotein E
20	1388.5	87.4	9360	24	AB013915	Human apolipoprotein E
21	1388.5	87.4	10716	24	AAD26034	Human apolipoprotein E
22	1383	87.0	3805	20	AA020524	Human apolipoprotein E
23	1383	87.0	3805	20	AA020526	Human apolipoprotein E
24	1361.5	85.7	10716	24	AAD26108	Human apolipoprotein E
25	1162.5	73.2	965	24	AAD32081	Human apolipoprotein E
26	1162.5	73.2	5617	24	AAD32077	Human apolipoprotein E
27	1162.5	73.2	6026	24	AAD32075	Human apolipoprotein E
28	1130	71.1	1126	19	AA028159	Human apolipoprotein E
29	977	61.5	936	15	AA069101	Human apolipoprotein E
30	977	61.5	936	17	AA018070	Human apolipoprotein E
31	976	61.4	660	18	AA069792	Human apolipoprotein E
32	939	59.1	597	17	AA018068	Human apolipoprotein E
33	930	58.5	597	15	AA060909	Human apolipoprotein E
34	879	55.3	1381	22	AA022673	Human apolipoprotein E
35	877.5	55.2	786	21	AA018114	Human apolipoprotein E
36	670	42.2	600	20	AA089595	Human apolipoprotein E
37	651	41.0	407	24	AB034238	Human apolipoprotein E
38	625	39.3	478	24	AB034238	Human apolipoprotein E
39	625	39.3	478	24	AB034238	Human apolipoprotein E
40	625	39.3	478	24	AB034238	Human apolipoprotein E
41	594	37.4	499	22	AA030349	Human apolipoprotein E
42	506	31.8	330	12	AA011980	Human apolipoprotein E
43	504	31.7	405	21	AA020139	Human apolipoprotein E
44	485	30.5	345	22	AA098479	Human apolipoprotein E
45	411	25.9	260	21	AA040342	Human apolipoprotein E

ALIGNMENTS

RESULT 1	
AAD22051	
ID	AAD22051 standard: DNA: 1156 BP.
AC	AAD22051:
XX	
AC	12-FEB-2002 (first entry)
XX	
DT	
XX	
DE	Human apolipoprotein E (apoE) allele, apoE2* DNA.
XX	
KW	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW	hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX	
OS	Homo sapiens.
XX	
FX	
FT	key
FT	CDS
FT	Location/Qualifiers
FT	61..1014
FT	/*tag= a

FT sig.peptide /product= "Human apoE allele, apoE2*" /tag= b 61..114
FT mat.peptide /tag= c 115..1011
FT /product= "Mature human apoE allele, apoE2*" /tag= c
PN W0200177136-A1.
XX 18-OCT-2001.
XX 06-APR-2001: 2001MO-US11358.
PF 06-APR-2000: 2000US-0544386.
PR 04-OCT-2000: 2000US-0679088.
PR 05-APR-2001: 2001US-0827854.
XX (KOSP-) KOS PHARM. INC.
PA (UYBO-) UNIV BOSTON.
PI Zannits VI, Kypreos KE;
XX P-PSDB; AAE13297.
DR WPI: 2002-010885/01.
XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridemia -
XX Claim 14; Page 82; 91pp; English.
XX The present sequence is a human apolipoprotein E (apoE)
CC allele, apoE2* DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other:

Alignment Scores:
Pred. No.: 3.49e-116 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-18 (1-317) x AAD22051 (1-1156)
QY 1 MetLysValLeuTrpAlaIleuValThrPheLeuAlaGlyCysGlnAlaIleVal 20
DB 61 ATAAAGGTTCTGGGGCTGGTGTGCTGCTGTCACATTCCTGGCAGATGCCAGGCGT 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGACCTCGCCAGACAGCCAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaIleuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGCACACGCTGGGAACTGGCACTGGGCTCTTTGGGATTACTGCGCTGGTGACAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTCTGACCAAGTGCAGAGAGAGCTGCTCAGCTCCAGCTCACCCAGAACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 301 CTGATGACAGACCATGAAAGAGTTGAAGGCTTACAATCGAATCGAGAGAACACTG 360

QY 101 ThrProValAlaGluGluTrpArgAlaIleuLeuSerLysGluLeuGlnAlaIleAla 120
DB 361 ACCCGGTGGCGAGAGAGACCGCGGCACGCTGTCCAGAGACTGCAGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
DB 421 GGGCTGGGCGGACATGGAGACGTGTGGGCGGCTGTGGTGAATACCGCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCATGCTGGCGCAGACACCGAGAGCTCGGGGTGGCTCCGCTCCACCTGCGCGG 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTGTGACCGCGCTCCCGCCGATGCCGATGACCTTGAGAAAGCGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLysSerAlaIleArgGluArgLeuGly 200
DB 601 CAGCGCGGGCGCGAGGGGCGCGAGCGGCGCTCAGCGCCATCCGCGAGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGGTGGAAACAGGGCGCGCTGGCGGCCGCCACGTGGGCTCCCTGGCGCGCAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGCGCCAGCGCTGGCGGCGAGCGGCTCGCGCGCGATGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGAGCCCGACCGCCCTGGACGAGTGAGAGACAGCAGGTGGAGGTGGCGGCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 841 CTGAGAGAGAGAGCCAGCAGAGATAGCCTGACGAGCGGAGCGCTCCAGGCGCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGGTTGAGCGCCCTGTGTGAAAGCATGACAGCGCCAGTGGGCGCGCTGTGGAGAG 960
QY 301 ValGlnAlaIleValGlyThrSerAlaIleProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGCGGTGGCACACAGCGCGCCCTGTGCCAGGACATAC 1011

RESULT 2
AAD26035
ID AAD26035 standard; cDNA; 954 BP.
XX
AC AAD26035;
XX
DT 26-MAR-2002 (first entry)
DE Human apolipoprotein E (APOE) cDNA.
XX
KW Human; antihaemic; neuroprotective; noctropic; genetic variant; APOE;
KW apolipoprotein E; haployping; familial dysbetalipoproteinemia; therapy;
KW genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
KW atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..954
FT /tag= a
FT /product= "Human APOE protein"
FT /replace (13, C)
FT /tag= a
FT /standard.name= "Single nucleotide polymorphism (SNP)"
FT /replace (31, G)
FT /tag= b
FT /standard.name= "Single nucleotide polymorphism (SNP)"
FT /replace (364, A)
FT /tag= c
FT variation

FT /standard_name="Single nucleotide polymorphism (SNP)"
FT replace (388; C)
FT /tag= d
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT replace (498; G)
FT /tag= e
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT replace (526; T)
FT /tag= f
FT /standard_name="Single nucleotide polymorphism (SNP)"
FT replace (622; A)
FT /tag= g
FT /standard_name="Single nucleotide polymorphism (SNP)"
PN WO200179234-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US12303.
XX 14-APR-2000; 2000US-197188P.
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Chou JY, Klem SE, Koshy B, Lee HH;
XX
XX MPI: 2002-075064/10.
XX P-PSDB: AAE15158.
XX
XX Genotyping human apolipoprotein gene of individual for determining
XX haplotype of individual, involves determining identity of nucleotide
XX pair at specific polymorphic sites for two copies of gene -
XX
XX Claim 26; Fig 2; 78pp: English.
XX
XX The patent discloses novel genetic variants of human apolipoprotein
XX E (APOE) gene. The invention also relates to compositions and methods
XX for haplotyping and/or genotyping the APOE gene. The haplotyping
XX methods of the invention are useful for improving the efficacy and
XX reliability of several steps in the discovery and development of
XX drugs for treating diseases associated with APOE activity, e.g.
XX familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
XX atherosclerosis, and Alzheimer's disease. They are useful to validate
XX APOE as a candidate agent for treating a specific condition or disease
XX predicted to be associated with APOE activity and in the design of
XX clinical trials of candidate drugs for treating a specific condition
XX or disease predicted to be associated with APOE activity. Genotyping
XX or haplotyping methods are useful to screen for compounds targeting
XX APOE to treat a specific condition or disease associated with APOE
XX activity. The present sequence is a cDNA encoding human APOE protein.
XX APOE gene is located on chromosome 19q13.2.
XX
SQ Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;
Alignment Scores:
Pred. No.: 2.47e-115
Score: 1577.00
Percent Similarity: 99.68%
Best Local Similarity: 99.68%
Query Match: 99.24%
DB: 24 Gaps: 0
US-09-827-854-18 (1-317) x AAD26035 (1-954)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 1 ATGAAAGTCTGTGGGCTCGTTGCTGCATCTCTGGAGATGCCAGGCCAAGGTG 60
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnGlnSer 40
DB 61 GAGCAAGCGGTGAGACAGCCGAGCCGAGCTCGCCAGCAGCCAGTGGCAGAGC 120
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

DB 121 GGCACAGCGCTGGACATGCGACTGGGCTCGCTTTGGAGTATCACTGGCGTGGCAGACA 180
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 181 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCACCTCCAGTCAACCGAGACTGAGAGCGG 240
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 241 CTGATGGACGACACCATGTAAGAGCTTGAAGGCTTACAAATCGGAAGTGGAGACACTG 300
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 301 ACCCGCGTGGCGAGAGAGACCGGGGACCGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 360
QY 121 ArgLeuGlyAlaAspMetLysAspValLysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 361 CGGCTGGCGCGACATGAGAGACGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 421 CAGGCGATGCTGGCCAGACACCGAGAGCTGCGGGTGGCTGCTGCTGCTGCTGCTGCTG 480
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 481 AAGCTCGTAAAGGGGCTCTCCGCGCATGCGCATGACCTGACAGAGCGCTGGCAGGTAC 540
QY 181 GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaThrGluArgLeuGly 200
DB 541 CAGGCGGGGGCGCGCGAGCGCGCGCGCTGCGCGCTGCGCGCATCCGCGAGCGCTGGGG 600
QY 201 ProLeuValGlnGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 601 CCCCTGTGGACAGGGCGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 661 CTACAGAGCGGGCGCGCGAGCGCTGCGCGCGAGCGCTGCGCGCGATGAGAGATGGGC 720
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 721 ACCCGGACCGCGCGCGCTGCGCGAGGTGAGAGACAGTGGCGGAGGTGCGCGCGCAAG 780
QY 261 LeuGlnGluGlnAlaGlnGlnLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 781 CTGGAGGACAGCGCCAGAGATACCTGCGAGGCCGAGGCCCTTCCAGGCCCGCTCAGAG 840
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 841 ACCTGTTGAGAGCCCTGCTGGAAGACATGACAGCGCCAGTGGGCGCGGCTGGTGGAGAG 900
QY 301 ValGlnAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 901 GTGCAGGCTGCGGTGGGCACAGCGCGCGCTGTGCTCCAGGACCAATCAGC 951
RESULT 3
ID AAN60409 standard; DNA; 1110 BP.
AC AAN60409;
XX
XX 01-JAN-1980 (first entry)
XX
XX Human apolipoprotein-E.
XX
XX Apolipoprotein-E; hyperlipidemia; atherosclerosis; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FH 15..968
XX CDS
XX FT /*tag= a
XX

PN A08547513-A.
 XX 24-APR-1986.
 PD 17-SEP-1985; 85AU-0047513.
 PF 11-JUN-1985; 85UP-0126989.
 PR (MITU) MITSUBISHI CHEM IND KK.
 PA WPI: 1986-150217/24.
 DR P-PSDB; AAP60507.
 XX
 PT New DNA sequence coding for human apolipoprotein-E - and
 PT expression vectors and transformed cells contg. it
 XX
 PS Disclosure; Fig 2; 45pp; English.
 CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.
 XX
 SQ Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.93e-115 Length: 1110
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 7 Gaps: 0
 US-09-827-854-18 (1-317) x AAN60409 (1-1110)
 QY 1 MetysValLeuThrPalaIaLeuLeuValThrPheLeuAaGlyCysGlnAlaLysVal 20
 Db 15 ATGAAGGTTCCTGTGGGTGGCTGCTGTCACATCTCTGGCAGAGATCCAGGCCAAGGTG 74
 QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
 Db 75 GAGCAACCGGTGGAGACAGACCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAAGC 134
 QY 41 GlnGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 135 GGCACGCGCTGGGAATGCACTGCTGCTGCTTTGGATTACTGCGCTGGGTGCAGACA 194
 QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 Db 195 CTGTCTGACGAGGTGCGAGGAGAGCTCTCAGCTCCAGGTCACCGAAGCTGAGGGCG 254
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 Db 255 CTGATGACGACGACCATGAAGAGCTTGAAGCCCTACAAATCGAAGCTGAGGAACACTG 314
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 Db 315 ACCCCGCTGGCGGAGGAGCGGGCGCAGCTGTCCAAAGAGTGCAGCGCGCGCAGGCG 374
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
 Db 375 CGGCTGGGGCCCGACATGAGAGAGCTGTGGCGCCCTGTGTGACATCCCGCGGAGAGTG 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 435 CAGGCCATGTCTCGGCCAGACACCGAGAGCTGGGGTGGCCCTCGCTCCACACTGCC 494
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 495 AAGCTGCGTAAAGCGGCTCTCGCGATGCCAGATGACCTGCAGAAAGCCCTGGCAGTATC 554
 QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

Db 555 CAGCCGGGGCCCGGAGGGCGCCGAGCGCGCTCAGCGCATCCCGAGCGCTGGG 614
 QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 615 CCCCTGTGGAAACAGGGCCCGGTGGCGGCCCACTGTGGGTCTCTGGCCGGCAGCCG 674
 QY 221 LeuGlnGlnArgAlaGlnAlaTyrGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 Db 675 CTACAGAGAGGGGCCAGGCTGTGGCGAGGCGCTGGCGCGCGGATGAGAGATGGCG 734
 QY 241 SerArgThrArgAspAlaGlyLeuAspGlyValLysGlnGlnValAlaGlnValArgAlaLys 260
 Db 735 AGCCGGACCCCGCCCGCTGGAGAGGTGAAGAGAGGTGGCGGAGGTGGCCAAAG 794
 QY 261 LeuGlnGlnGlnAlaGlnGlnGlnThrSerAlaAlaProValProSerAspAsnHis 280
 Db 795 CTGGAGAGACGAGCCCGACAGATACGCTTCAGAGCCAGAGCCCTTCAGAGCCGCTCAAG 854
 QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 Db 855 AGCTGTTTCAGGCCCTGTGTGAAGACATGCAGCGCCAGTGGCGGGCTGTGGAGAG 914
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 915 GTGCAGCTGCGCTGGGACACGCGCCGCTGTGTGCCAGCGACAATCAC 965
 RESULT 4
 ABA83113
 ID ABA83113 standard; DNA; 1147 BP.
 AC ABA83113;
 XX 08-FEB-2002 (first entry)
 DT
 XX
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
 XX
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumor; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumor;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumor; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumor; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200175177-A2.
 XX
 PD 11-OCT-2001.
 PF 03-APR-2001; 2001WO-US10947.
 PR 03-APR-2000; 2000US-194336P.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 DR WPI: 2001-626450/72.
 DR P-PSDB; ABB50287.
 XX
 PT detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 gene -
 PS Claim 23; Page 105-106; 140pp; English.
 CC The invention relates to methods for diagnosing and prognosing ovarian

CC tumours in an individual via the detection and measurement of the
 CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 CC ABA83181 and ABA83183). The methods of the invention are useful for
 CC detecting an ovarian tumour in a patient, for identifying an individual
 CC at increased risk for developing ovarian cancer, in prognostic tests for
 CC assessing the relative severity of ovarian cancer, in tests for
 CC monitoring a patient in remission from ovarian cancer and in tests for
 CC monitoring disease status in a patient being treated for ovarian cancer.
 CC The methods can additionally be used to identify a particular tumour as
 CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 CC mucinous cystadenoma, borderline mucinous tumour, mucinous
 CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 CC tumour. The ovarian tumour marker genes of the invention were identified
 CC using SAGE (Serial analysis of gene expression) and were found to be
 CC overexpressed in a broad variety of ovarian epithelial tumour cells
 CC relative to normal ovarian epithelial cells. The marker genes are
 CC implicated in immune response pathways, in the regulation of cell
 CC proliferation and in protein folding, and many of these are membrane-
 CC localised or secreted. In addition to their use as diagnostic and
 CC prognostic markers, the ovarian tumour marker genes or their encoded
 CC proteins may be used as therapeutic targets for the treatment and
 CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 CC the invention.

XX Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 3,04e-115 Length: 1147
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.688 Conservative: 0
 Best Local Similarity: 99.688 Mismatches: 1
 Query Match: 99.24% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-18 (1-317) x ABA83113 (1-1147)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 46 ATGAAGTCTCTGGGCTCGTGTGTCGTCACATCTCTGGACAGATGCCAGGCAAGGTC 105
 QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 106 GACCAAGCGGTGGAGACAGACCGGAGCCGAGACTCGCCAGCAGACGAGTCGCGAGAC 165
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 166 GCGCAGCGCTGGGAACCTGGCCTCTTTTGGATTACCTGGCTGGGTCAGACA 225
 QY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 226 CTGTCTGACAGAGTCAGAGAGCTGCACGCTCCAGATCCAGACCAAGAACTGAGGGG 285
 QY 81 LeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 286 CTGATGGAGAGACCATYGAAGAGTGAAGGCTTCAAAATCGGAATGGAGGAACAACG 345
 QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 346 ACCCGGTGGCGAGAGACCGCGGCTCTCCAAAGAGCTCCAGCGCGCGCGAGCC 405
 QY 121 ArgLeuGlyAlaAspMetCysValLysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 406 CGGCTGGCGCGACATGAGAGCGTGTGCGCGCTGTGCAGTAACCGCGCGAGGTG 465
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 466 CAGGCGCATCTCGGCGAGACACCGGAGAGCTGGGGTGGCTCCCTCCACCTCGCGC 525
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 526 AAGCTCGTAAGGGGCTCTCCGCGATGCGGATGACCTGCAGAAAGCGCTGGCAGTGTAC 585
 QY 181 GlnAlaGlyAlaArgGluGluValGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 586 CAGGCGGGGGCCCGCAGAGGGCGCGGCGGCTTCAGCGCATCCGCGAGCGCTGGGG 645
 QY 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 646 CCCCTGTGGAAACAGAGGCGCGCTGGCGCCCTCTGTGGCTCTCTGGCGCGCAGCCG 705
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 706 CTACAGAGACGGGCGCCAGGCTGTGGCGAGCGGCTCGCGCGCGGATGGAGAGATGGGC 765
 QY 241 SerArgTrpArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 766 AGCCGAGACCCGCGACCGCTGTGACGAGGTGAAGGACAGGTGGCGAGGTGGCCCAAG 825
 QY 261 LeuGlnGluGlnAlaGlnGlnLeuArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 826 CTGGAGGAGACAGGCCAGCAGATACGCTGTGAGGCCGAGGCTTCCAGGCGCGCTCAAG 885
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 886 AGCTGTGTGAGCGCTGTGTGGAAGACATGACAGCGCAGTGGCGCGGCTGTGTGAGAA 945
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 946 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGTCCAGACCAATAC 996
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 5
 AAF84315
 ID AAF84315 standard; cDNA; 1156 BP.
 XX
 XX AAF84315;
 AC
 XX
 XX 21-JUN-2001 (first entry)
 DT
 XX Human ApoE3 coding sequence.
 DE
 XX Human; ApoE3; Alzheimer's disease; arteriosclerosis; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH 61..1014
 FT /*tag= a
 FT /product= "Human ApoE3"
 FT
 XX JP2001017028-A.
 PN
 XX
 XX 23-JAN-2001.
 PD
 XX 28-APR-2000; 2000JP-0128919.
 PF
 XX 06-MAR-1999; 99JP-0125647.
 PR
 XX (MITU) MITSUBISHI CHEM CORP.
 PA
 XX
 XX WPI: 2001-285406/30.
 DR P-PSDB; AAB80997.
 DR
 XX
 XX New apoE humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 PT
 PS Disclosure: Page 13-14; 22pp; Japanese.
 CC
 CC The present invention relates to an ApoE humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE3, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing

US-09-827-854-18 (1-317) x AAD22048 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGGCTGTGTCACATTCTTGCGAGATGCGCAGGCAAGGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuAlaLeuAlaArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 181 GGCCAGCGCTGGAGACTGGACACTGGGTGCTTTGGAGATTACTGGTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 241 CTGTCTGACAGAGTGCAGAGAGAGTGTCTACACTCCAGCTCACCAGCACTGAGGCGG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGGACGACACCATGAGAGATTGAGAGCTTACAAATCGGAATGAGAGACAACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGAGCGGCGACGCGCTGCCAAGAGAGCTGACAGCGCGCAGAGCC 420
 QY 121 ArgLeuGlyAlaAspMetCysAlaGlyArgLeuValGlnTrpArgGlyGluVal 140
 DB 421 CGGTGGGGCGGACATGAGAGAGCTGTGCGGCGCTGTGTACAGTACCGCGGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerLysLeuArg 160
 DB 481 CAGGCGATGCTGGCCAGAGACCGAGGAGCTGCGGGTGGCTGCTGCCCTCCACCTGCGCC 540
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTrp 180
 DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACCTCAAGACGCTGCGCAGTGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 601 CAGGCGGGGGCGGCGAGGCGCGCCAGCGCGCTCAGCGCCATCGCGAGCGCCTGGGG 660
 QY 201 ProLeuValGlnGlnGlyValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 661 CCCCTGTGTGAGACAGAGGCGCGTGGGCGCGCCACTGTGTGCTCCCTGCGCGCAGCGC 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
 DB 721 CTACAGAGAGCGGCGCGAGGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGCGC 780
 QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 781 ACCCGGACCGCGGACCGCTGTGACGAGGTGAAGGAGCAGGTGCGAGTGGCGCGCCAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGGAGAGACAGCGCCAGCAGATACGCTGCGAGCGCGAGGCTTCCAGGCGCGGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 901 ACCTGTTGAGACCGCGCTGGTGAAGCATGACAGCGCCAGTGGCGCGGTGTGTGAGAAAG 960
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGACAGGCTGCGGTGGGACACGCGCCGCTGTGTGCCAGGACAAATCAC 1011
 RESULT 7
 AAD22052
 ID AAD22052 standard: DNA; 1156 BP
 AC
 AC AAD22052;
 XX

DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 KW Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 61..1014
 FT /*tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT mat_peptide 115..1011
 FT /*tag= b
 FT /*tag= c
 FT /*product= "Mature human apoE allele, apoE2**"
 XX
 PN W0200177136-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-US11358.
 XX
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kyriacos KE;
 DR WPI; 2002-010885/01.
 DR P-PSDB; AAE13298.
 XX
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia
 XX
 PS Claim 14; Page 83; 91pp; English.
 XX
 CC The present sequence is a human apolipoprotein E (apoE)
 CC allele, apoE2** DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 SQ Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 6,33e-115
 Score: 1573.00 Length: 1156
 Percent Similarity: 99.68% Matches: 315
 Best Local Similarity: 99.37% Mismatches: 1
 Query Match: 98.99% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-18 (1-317) x AAD22052 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGTGGGCTGGCTGTGTCACATTCTTGCGAGATGCGCAGGCAAGGTG 120
 QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC 180

QY 41 GlytlnArgrTPgluLeuAlaLeuGlyArgrPhetrrpAsprYrleuArgrTPrValGlnThr 60
DB 181 GGCCAGCGCTGGAGACCTGGCTGCTTTGGATTTCCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGCTGTGAGCAGGTGTGACAGAGAGACTGCTCAGCTCCAGGTCCACCGAAGACTAGGGCG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGAGCAGACCATGAAGAGAGTGAAGCCCTACAAATCGAAGTGAAGACAACACTG 360
QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGGGAGAGAGACGGGGCAGCGCTGTCCAAAGAGTGCAGGGCGGCAGAGCC 420
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140
DB 421 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCGCTCGTGCAGTACCGCGGGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGGCCAGAGACCCAGAGAGCTGGGGTGGCTCCCTCCACCTGCGC 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGTGCTGAGCGGCTCTCCGATGCCGATACCTGCGAAGCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
DB 601 CAGCGCGGGGGCGGCGAGGGCGCGGCGCTCAGCGCCATCCCGAGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
DB 661 CCCCTGTGTGAACAGGCGCGCTGCGGGCGCCACTGTGGGCTCCCTGGCGGCGCAGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaThrGlnGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 721 CTACAGAGAGCGGGCCAGGCTGTGGGGAGCGGCTGGCGCGGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 781 AGCCGGAGCCCGGACCCGCTGTGAGAGGTGAGAGAGCAAGGTGGCGGAGGTGCGGCCAAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGGAGAGAGCGGCCAGAGATAGCGCTGCGAGGCGCTTCAGGCCCGGCTCAAG 900
QY 281 SerTyrPheGlnProLeuValGlnAspMetGlnArgGlnTyrPheGlnValGlnLys 300
DB 901 AGCTGTTTCAGACCCCTGTGTGGAAGACATGACAGCGCAGTGGCGCGGTGTGTGAGAG 960
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGAGAGCTGCGCTGGGCGACAGCGCGCCCTGTGCCCCAGCCACATCTAC 1011
RESULT 8
AAf84314
ID AAf84314 standard; cDNA; 1156 BP.
XX
AC AAf84314:
XX
DT 21-JUN-2001 (first entry)
XX
DE Human ApoE2 coding sequence.
XX
KW Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX
OS Homo sapiens.
XX
FH Key 61..1014 Location/Qualifiers
FT CDS 61..1014
FT /tag= a

FT /product= "Human ApoE2"
PN JP2001017028-A.
XX
XX 23-JAN-2001.
PD
XX 28-APR-2000; 2000JP-0128919.
PF
XX 06-MAY-1999; 99JP-0125647.
PR
XX (MITU) MITSUBISHI CHEM CORP.
PA
XX WPI: 2001-285406/30.
DR
DR P-PSDB; AAB80996.
XX
XX New apoE humanized mammalian cell useful for screening for agents
PT useful for treating or preventing Alzheimer's disease and
PT arteriosclerosis -
XX
XX Disclosure; Page 11-12; 22pp; Japanese.
XX
XX The present invention relates to an ApoE humanised mammalian cell. The
CC present sequence is the coding sequence for human ApoE2, which was used
CC in the method of the present invention. The ApoE humanised mammalian cell
CC can be used for screening for agents useful for treating or preventing
CC Alzheimer's disease and arteriosclerosis.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 1 31e-114 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.74% Indels: 0
DB: Gaps: 0
US-09-827-854-18 (1-317) x AAf84314 (1-1156)
QY 1 MetLysValLeuTrrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAGAGTTCTGTGGGTGGCTGTGCTGTGCTACATTCCTGCGAGAGTCCAGGCCAGAGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTyrGlnSer 40
DB 121 GAGCAAGCGGTGAGAGACAGAGCGGAGCCGACCTCGCCAGCAGACCGCAGTGGCAGAGC 180
QY 41 GlytlnArgrTPgluLeuAlaLeuGlyArgrPhetrrpAsprYrleuArgrTPrValGlnThr 60
DB 181 GGCCAGCGCTGGAGACCTGGCTGCTTTGGATTTCCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGCTGTGAGCAGGTGTGACAGAGAGCTGCTCAGCTCCAGGTCCACCGAAGACTAGGGCG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGAGCAGACCATGAAGAGAGTGAAGCCCTACAAATCGAAGTGAAGACAACACTG 360
QY 101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGGGAGAGAGACGGGGCAGCGCTGTCCAAAGAGTGCAGGGCGGCAGAGCC 420
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140
DB 421 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCGCTCGTGCAGTACCGCGGGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGGCCAGAGACCCAGAGAGCTGGGGTGGCTCCCTCCACCTGCGC 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

DB 541 AAGCTCCGTAAGCGCTCCCTCCGCGATGCCGATGACCTGACAGAAAGTCCCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGGCGGGGGCCCCGAGGGGCGCCGAGCGGCGCTGAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGluGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 661 CCCCTGTGTGAACAGAGGCGCGCTGGCGGCGGACACTGTGGCTCCCTGGCCGCGCAGCGG 720
QY 221 LeuGluGluArgAlaGluAlaThrPglGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
DB 721 CTACAGGAGCGGGCGCCAGCGCTGTGGCGCGAGCGCGCTGCGCGCGATGAGAGATGGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLeuGlyGluGluValAlaGluValArgAlaGly 260
DB 781 AGCCGAGACCCGCGACCGCTGGACGAGGTGAAGAGACAGGTGGAGGTGGCGCGCAAG 840
QY 261 LeuGluGluGluAlaGluGluGluIleArgLeuGluAlaGluAlaPheGluAlaArgLeuGly 280
DB 841 CTGAGAGAGACAGGCGCCAGAGATAGCCCTGCAAGCGCGGCTTCAGGCGCGCGCTCAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGluArgGluTrpAlaGlyLeuValGluGly 300
DB 901 ACCTGTGTGAGCCCTGTGTGAGAGACATGACAGCGCCAGTGGGCGCGCTGGTGAGAG 960
QY 301 ValGluAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGCGCTGGGACACAGCGCGCCCTGTGCTCCAGCGACATCATC 1011
RESULT 9
AAD22049
ID AAD22049 standard; DNA; 1156 BP.
AC AAD22049;
DT 12-FEB-2002 (first entry)
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
XX Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;
KM hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT CDS 61..1014
FT sig_peptide /tag= a
FT mat_peptide /tag= b
FT /tag= C
FT /Product= "Mature human apoE isoprotein, apoE2"
XX
XX WO200177136-A1.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-US11358.
XX
XX 06-APR-2000; 2000US-0544386.
XX 04-OCT-2000; 2000US-0679088.
XX 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
XX (UYBO-) UNITV BOSTON.
XX
XX Zannis VI, Kypreos KE;
XX
XX WPI: 2002-010885/01.
XX P-PSDB: AAE13295.
XX

PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia -
XX
XX Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 1..31e-114 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.74% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-18 (1-317) x AAD22049 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 61 ATGAAGGTTCTGTGGGCTCGTTGCTGTCACATTCCTGCGAGATGCCAGGCCAAGCTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGluGluThrGluTrpGlnSer 40
DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCACAGACCAAGTGGCAGAC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 181 GGCACAGCGTGGGAACCTGCGACTGGGTGCTTTGGGATTACCTGCGTGGGTGAGACA 240
QY 61 LeuSerGluGluValGluGluGluLeuLeuSerSerGlnValThrGluLeuArgAla 80
DB 241 CTGTGTGACAGAGTGCAGAGAGGCTGTCAGCTCCAGCTCACCCAGCACTGAGGCGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlu 100
DB 301 CTGATGACGAGACCATGAGAGGTTGAGAGCTTCAAAATCGAATGAGGACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGCGGAGAGAGCGGGGCGAGGCTGCCAAGAGCTGCGAGCGCGCGAGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGluTrpArgGlyGluVal 140
DB 421 CGGCTGGGCGCGACATGAGAGACGTGCGCGCGCTGTGTCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGGCGAGAGACCGAGAGCTGCGGGTGGCTGCTCCACCTGGCGC 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGluLysArgLeuAlaValTrp 180
DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGGCGATGACCTCAGAAAGTCCCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluValArgGlyGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGGCGGGGGCCCCGAGGGGCGCCGAGCGCGCTCAGCGCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGluGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 661 CCCCTGTGTGAACAGAGGCGCGCTGGCGGCGGACACTGTGGCTCCCTGGCCGCGCAGCGG 720
QY 221 LeuGluGluArgAlaGluAlaThrPglGlyGluArgLeuArgAlaArgMetGluGluMetGly 240

Df			
Df	721	CTACAGGAGCGGGCCACGAGCCTGGGGCGAGCGGTGGCGCGCGCGATGAGAGATGGC	780
Oy	241	SerArYthrArGaSPaRgLeuAsPGLValLysGlucInValAlaIgluValArGaLaLys	260
Df	781	AGCCGACCACCCGCAGCCCGCTGCAGAGGTGAAGAGCAGGTCGGCGAGTGCGCCAAG	840
Oy	261	LengluJugluJnAlaglnGlnIlleargLeuGlnlAlaIgluAlaPhcgnAlaIarglyLeLys	280
Df	841	CTGGAGGAGCAGGGCCACCAATACGCTGCAGGCGCAGGCGCTTCCAGGCCCCCTTAAG	900
Oy	281	SerTrPPheGluPProLeuValGIaSpwmetLArgGIntPrPaIagLyLeuValGIuLys	300
Df	901	AGCTGTTCTGAGACCCCTTGTTGTAAGACTGACAGGCCACAGTGGCGCGGCTGTGGAGAG	960
Oy	301	VaIglnAlaAlaValaIGlyThrSerAlaAlaProvalProSerAspaSnHIS	317
Df	961	GTCGAGGCTGCCTGGGGCAGCGCGCGCCCTGTGCCAGCACCATTCAC	1011
 RESULT 10 AAr06957 ID AAr06957 standard; cDNA to mRNA; 1157 BP. XX XX AC AAr06957; XX DT 19-JUN-1996 (first entry) XX DE Human apolipoprotein-E (APOE) cDNA. XX KM Recombinant; human; apolipoprotein-E; APOE; insect cells; larva; KW Manduca sexta; Autographica californica nuclear polyhedrosis virus; XX haemolymph; lipid complex; biologically active; ds. XX OS Homo sapIens. XX XX FH Key location/Qualifiers FT misc_feature 62..119 FT note= "/*tag= a FT note= \"misc_signal\" FT CDS 62..1015 FT /*tag= b XX XX US5472858-A. XX PD 05-DEC-1995. XX PF 04-JUN-1991; 91US-0709949. XX PR 04-JUN-1991; 91US-0709949. XX PA XX (WISC) WISCONSIN ALUMNI RES FOUND. XX PA PI Attie AD, Beckage NE, Gretch DG, Sturley SL; XX WP: 1996-029812/03. DR P-PSDB; AAR86791. XX PT Prodn. of recombinant apo:lipo:protein E In insects - by infecting PT Manduca sexta larvae with recombinant Autographica californica PT nuclear polyhedrosis baculovirus vector. XX XX Disclosure; Columns 11-14; 10pp; English.			
CC	Recombinant human apolipoprotein-E (APOE) (AAR86791) can be produced		
CC	by preparing a genetic construct (contg. an ApoE-encoding sequence,		
CC	e.g. AAr06957, and flanking regulatory sequences enabling the protein		
CC	to be expressed in insect cells), which is then introduced into a		
CC	Manduca sexta larva (using a recombinant Autographica californica		
CC	nuclear polyhedrosis virus) and recovering the protein from the		
CC	haemolymph of the larval host. The ApoE produced is in a form		
CC	sufficiently complexed with lipids to be biologically active, which		
CC	cannot be achieved in insect cell cultures, and can therefore be		
CC	used in therapeutic applications		

[illegible]

XX WO200212440-A2.
 PN 14-FEB-2002.
 XX 07-AUG-2001; 2001WO-US24708.
 PF 07-AUG-2000; 2000US-223323P.
 XX 05-JUN-2001; 2001US-0873319.
 XX (GENE-) GENE LOGIC INC.
 PA (MISB) JAPAN TOBACCO INC.
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 DR WPI; 2002-257476/30.
 XX
 PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells -
 XX
 PS Disclosure; Page 239-240; 444pp; English.
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABR64106-ABR64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.
 CC
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 1.57e-114 Length: 1157
 Score: 1568.00 Matches: 314
 Percent Similarity: 99.05% Conservative: 0
 Best Local Similarity: 99.05% Mismatches: 3
 Query Match: 98.68% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-18 (1-317) x ABR64514 (1-1157)
 QY 1 MettysValleuTrpAlaIalaLeuValThrpheLeuAlaIagIyCysGlnAlaIysVal 20
 DB 62 ATGAAGGTCTGTGGGTGGCTGTCTGTCTGTCATCTCTGGAGGATCCCAAGGCAAGTGTG 121
 QY 21 GlnGlnAlaValaGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer 40
 DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGTGGCCAGCAGACCGATGGCAAGC 181
 QY 41 GlyGlnAArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTyrLeuAArgTTPValGlnThr 60
 DB 182 GGCCAGGCGTGGGAACTGGCACTGGCTTTGGGATTAACGCGCTGGGTGCACAGACA 241
 QY 61 LeuSerGlnGlnValaGlnGlnGluLeuLeuSerGlnValThrpGlnGluLeuAArgAla 80
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 242 CTGTCTGAGCAGGTGCAGAGAGACCTGCTCAGCTCCCAAGTCCACCAAGAACTGAGGCGC 301
 QY LeuMetAspGluThrMetLysGluLeuIysAlaTyrLysSerGluLeuGlnGluLeu 100
 DB 302 CTGATGACGACGACCATGAAGAGACTTGAAGCCCTACAAATCGAACTGGAGAACTG 361
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 DB 362 ACCCGGTAGCGAGAGACCGCGGACAGCGGTGTCCAAAGAGCTGCAGACGCGCAGAGCC 421
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGluArgLeuValGlnTyrArgGlyGluVal 140
 DB 422 CGGCTGGGCGCGACATGAGAGACGTGTGCGCCGCTGTGTCAGTACCCTGGGAGAGTGG 481
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCCATGCTCGGSCCAGAGACCGCAGAGAGCTGGGGTGGCCCTCCGCTCCACCTGGCC 541
 QY 161 LysLeuGlnLysArgLeuLeuArgAlaAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAGACCGCTGGCAGTGTAC 601
 QY 181 GlnAlaGlnAlaAArgGluGluGlnAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 602 CAGCGCGGGGCGCGGAGGGCGCGAGCGCGCTCAGCGCCATCCCGCAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 662 CCCCTGGTGAACAGGGGCCGCTGGCGGCCGACCTGTGGCTCTCCCTGGCGGCGACCGC 721
 QY 221 LeuGlnGluArgAlaGlnAlaTTPGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 DB 722 CTACAGAGAGCGGGCGCGAGCTGGGCGAGCGCGCTGCGCGCCGCGAGTGAAGATGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValaGlnGluValArgAlaLys 260
 DB 782 AGTGGACCCGACGCCCTGTGAGAGAGTGAAGAGCAGAGTGGCGAGGTGGCGCCAGAG 841
 QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGGAGAGAGCAGGCCACAGATACGCTGACGAGCCGAGGCGCTTCCAGGCGCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnAArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AGCTGGTTCAGACCCCTGGGTGAAGACATCAGAGCGCAGTGGCGGCTGGTGAAGAG 961
 QY 301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 962 GTGCAGGCTCGCGTGGGACACAGCGCCGCCCTGTGCCACGACATTCAC 1012
 RESULT 13
 ABL65450
 ID ABL65450 standard; DNA; 1157 BP.
 XX
 AC ABL65450;
 XX
 XX 15-MAY-2002 (first entry)
 DE Lung cancer related gene sequence SEQ ID NO:3787.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; aniltumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX

PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVALON PHARM.
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX MPI: 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1: SEQ ID 3787; 44pp: English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (1) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 1,57e-114 Length: 1157
Score: 1568.00 Matches: 314
Percent Similarity: 99.05% Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.68% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-18 (1-317) x ABL65450 (1-1157)
QY 1 MetLysValLeuThrPalaIalaLeuValThrPheLeuAlaGlyCysGlnAlaTyrVal 20
DB 62 ATGAAGTTCTGTGGGCTCGCTTGGCTGTCACATTCCTCGCAGAGTCCAGGCAAGGTG 121
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGCCGAGCCGAGCTGCGCAGACAGACAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GCCACAGCTGGGAAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 241
QY 61 LeuSerGlnGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 242 CTGTCTGACAGAGTGGAGAGAGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 301
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
DB 302 CTGATGACGAGACCATGAAGAGTGAAGGCTTACCAATCGAATCGAATCGAATCGAATCG 361
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 362 ACCCGGTAGCGAGAGAGAGCGGGGCGGCTGTCAGAGAGCTGAGAGCGGGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetLysAspValLysGlnArgLeuValGlnThrArgGlnGlnVal 140
DB 422 CGGCTGGGCGGACATGAGAGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGGCGCAGAGACCGAGAGAGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAAGCGGCTCCCGCGATCCCGATGACCTCCAGAGAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGlnAlaArgGlnGlnValArgGlnGlnLysLeuSerAlaIleArgGlnArgLeuGly 200
DB 602 CAGGCGGGGGCGGCGAGGGGGCGGAGCGGCGGCTTCAGGCGCATCCGAGCGGCTGGGG 661
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCGTGGTGAACAGAGCGCGGTGGCGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 721
QY 221 LeuGlnGlnArgAlaGlnAlaThrProGlnArgLeuArgAlaArgMetGlnGlnGly 240
DB 722 CTACAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaTyr 260
DB 782 AATCGGACCGCGGACCGCTGGACAGAGTGAAGAGAGAGAGTGGAGAGTGGCGGCGGCAAG 841
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGAGGAGACAGCGGCGGAGATGAGCTGAGGCGGAGAGCTTCACAGGCGCGGCTCAAG 901
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlnGlnValGlnLys 300
DB 902 ACCTGTTTCAGACCCCTGGTGAAGACATGAGCGGCGGCTGGCTGGTGAAGAG 961

PR 06-MAY-1999: 99JP-0125647.
 XX
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 DR WPI: 2001-285406/30.
 DR P-PSDB: AAB80998.
 PT New apoe humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 XX
 PS Disclosure: Page 16-17; 22pp; Japanese.
 XX
 CC The present invention relates to an Apoe humanised mammalian cell. The
 CC present sequence is the coding sequence for human Apoe4, which was used
 CC in the method of the present invention. The Apoe humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX
 SQ Sequence 1156 BP; 208 A; 369 C; 432 G; 147 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.7e-114 Length: 1156
 Score: 1565.00 Matches: 315
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 2
 Query Match: 98.49% Indels: 0
 DB: 22 Gaps: 0
 US-09-827-854-18 (1-317) x AAF84316 (1-1156)
 QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyGlnAlaLysVal 20
 DB 61 ATGAAGGTTCTGGGGCTGGCTGGTCACTTCTGCGAGATGCCAGGCCAAGGTG 120
 QY 21 GUGGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnSer 40
 DB 121 GAGCAGACGGGTGAGACAGAGCCGCGAGCTGCCAGACAGACAGTGGCAGACC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCACACGGCTGGGAACTGCGACCTGCTTTGGATTACCTGCGTGGGTGCAGACA 240
 QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCACCGAGAACTGAGAGCG 300
 QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
 DB 301 CTGATGAGACGACCATGAAGAGTGAAGGCTTACAATCGGAAGTGGAGAACACTG 360
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCCGGTGGCGAGAGAGCGGGCGCTGTCCAAAGAGCTGCAGGCGGGCGAGGCC 420
 QY 121 ArgLeuGlnValAspMetGlnAspValLysGlyArgLeuValGlnTyrArgGlyVal 140
 DB 421 CGGCTGGCGCGACATGAGAGAGCTGCGCGCGCTGTGTCAGTACCGCGGAGGTG 480
 QY 141 GlnAlaMetLeuGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCCATGCTGGCCAGAGACCGAGAGCTGCGGGTGGCTGCCCTCCACCTGCGG 540
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AAGCTGCTGAAGCGCTCTCCGCGATGCGGATGACCTGCAGAGCCCTGGCAGTGTAC 600
 QY 181 GlnAlaGlnValArgGlnGlnValArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 601 CAGGCCGGGGCCCGCAGAGGCGCGGCGCTCAGCGCCATCCGCAAGCGCTGGGG 660
 QY 201 ProLeuValGlnGlnLysArgValArgAlaIleThrValGlySerLeuAlaGlnPro 220

DB 661 CCCCTGTGGAACAGAGGCGCGTGGCGGCCGCACTGTGGCTCCCTGGCCGCGCACGCCG 720
 QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 DB 721 CTACAGAGCGGGCCAGGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
 DB 781 AGCCGAGCCCGCAGCCGCTGGAGCGAGGTGAAGAGACAGTGGCGAGGTGGCGCCAG 840
 QY 261 LeuGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 DB 841 CTGAGAGACAGGCGCCAGCAGATACGCTGACAGGCCGAGGCTTCCAGGCCCGCTCAAG 900
 QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
 DB 901 AGCTGGTTGAGCCCTGTGTGAAGCATGACAGCGGCACTGGGCGCGGCTGTGTGAAG 960
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 961 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGTGCCACAGCAATCAC 1011

Search completed: March 14, 2003, 12:43:20
 Job time : 183.728 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compen Ltd.

OW protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-18

Sequence: 1 MKVIMALLVTFLLAGCGQAKV.....VEKVOAAVTSAPVPSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frimex.p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/rnat_11032003.101611.27510/app.query.fasta_1.3576
-DB=Issued Patents.NA -OPMT=fastap -SUPER=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.@CGL_1.1.224.@rnat_11032003.101611.27510 -ICPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEGLSCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*

1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1568	98.7	1157	1	US-07-709-949-1
2	1130	71.1	1126	4	US-08-949-155-5
3	1130	71.1	1126	4	US-09-819-964-5
4	994	62.6	4267	4	US-08-949-155-51
5	994	62.6	4267	4	US-09-819-964-51
6	976	61.4	660	2	US-08-726-306A-28
7	503	31.7	330	1	US-07-849-389-6
8	366	23.0	252	3	US-08-617-256-24
9	366	23.0	252	4	US-09-287-141-24
10	366	23.0	252	4	US-09-431-613-24
11	366	23.0	252	4	US-09-504-245-24
12	366	23.0	252	4	US-09-287-682-24

13	366	23.0	252	4	US-09-287-679-24	Sequence 24, Appl
14	366	23.0	252	4	US-09-397-766-24	Sequence 24, Appl
15	366	23.0	252	4	US-09-287-681-24	Sequence 24, Appl
16	366	23.0	252	4	US-09-495-444-24	Sequence 24, Appl
17	177.5	11.2	842	4	US-08-952-796-1	Sequence 1, Appl
18	174.5	11.0	801	1	US-07-959-946-4	Sequence 4, Appl
19	174.5	11.0	801	1	US-08-333-577-4	Sequence 4, Appl
20	174.5	11.0	801	5	PCT-US92-08634-4	Sequence 4, Appl
21	168	10.6	5661	4	US-08-938-105-2	Sequence 2, Appl
22	157.5	9.9	863	1	US-08-448-606-5	Sequence 5, Appl
23	154.5	9.7	863	1	US-08-448-606-7	Sequence 7, Appl
24	153	9.6	1879	4	US-09-750-580-2	Sequence 2, Appl
25	150	9.4	3256	2	US-08-968-751-3	Sequence 3, Appl
26	147.5	9.3	603	4	US-08-952-796-14	Sequence 14, Appl
27	147	9.3	8789	1	US-08-328-254-5	Sequence 2, Appl
28	147	9.3	10136	1	US-08-353-700-2	Sequence 2, Appl
29	147	9.3	10136	5	PCT-US95-16216-2	Sequence 2, Appl
30	146	9.2	4868	1	US-08-139-937-12	Sequence 12, Appl
31	146	9.2	4868	5	PCT-US93-11310-12	Sequence 12, Appl
32	140.5	8.8	1960	2	US-08-533-306A-1	Sequence 1, Appl
33	140.5	8.8	1960	2	US-08-742-923A-1	Sequence 1, Appl
34	140.5	8.8	11236	1	US-07-853-913-1	Sequence 1, Appl
35	139.5	8.8	4852	1	US-07-853-913-3	Sequence 3, Appl
36	139	8.7	516	4	US-09-183-861-34	Sequence 34, Appl
37	139	8.7	516	4	US-09-022-765-34	Sequence 34, Appl
38	137	8.6	6706	1	US-08-195-487-3	Sequence 3, Appl
39	137	8.6	6306	5	PCT-US93-06160-3	Sequence 3, Appl
40	136.5	8.6	3319	1	US-08-006-676B-2	Sequence 2, Appl
41	136.5	8.6	3319	1	US-08-282-845-1	Sequence 1, Appl
42	136.5	8.6	3319	2	US-08-428-414A-4	Sequence 4, Appl
43	136.5	8.6	3319	5	PCT-US94-00324-2	Sequence 2, Appl
44	136.5	8.6	6306	1	US-08-466-390-3	Sequence 3, Appl
45	136.5	8.6	6306	1	US-08-470-950-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
; Sequence 1, Application US/07709949
; Patent No. 5472858
; GENERAL INFORMATION:
; APPLICANT: Attie, Daniel G
; APPLICANT: Sturley, Stephen L
; APPLICANT: Beckage, Nancy E
; TITLE OF INVENTION: Production of Recombinant Proteins in
; TITLE OF INVENTION: Insect Larvae
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: P.O. Box 2113
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/709,949
; FILING DATE: 19910604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 9629691801
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 1:

QY	1	MettysValleuTPPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	62	ATGAAGGTTCTGTGGCGCTGCTGTGTCATCTTCCGTGGCAGATGCCAGGCCAAGTGC	122
QY	21	GIuGlnAlaValGIuThrGIuProGIuLeuAArgGlnGlnThrGIuTPRGIuSer	40
Db	122	GAGCAAGCGGTGGAGACAGACCAGGAGCCGAGCTGGCGCCAGCAGACGAATGGCAGAGC	189
QY	41	GIyGlnAArgTPRGIuLeuAlaAlaLeuGlyAArgPheTPRAspTyrLeuAArgTPRValGlnThr	60
Db	182	GGCCAGCGCTGGGAACATGGCACTGGGTGCGCTTTGGGATPRACCTGGCGCTGGGTGCACACA	244
QY	61	LeuSerGIuGlnValGlnGIuLeuLeuSerSerGlnValTPRnglnGIuLeuAArgAla	80
Db	242	CTGTCTGAGCAGGTGGCAGAGAGACTGCTACCTCCCAAGTCACCCAAAGCATGAAGCGCG	304
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGlnGIuGlnLeu	100
Db	302	CTGATGGACGAGACCATGAAGAGATTGAAGCCCTCAAAATGGAACTGGAGAACCAACTG	364
QY	101	ThrTyrValAlaGIuGIuThrArgAlaAArgLeuSerLysGIuLeuGlnAlaAlaGlnAla	120
Db	362	ACCCGGGTAGCGGAGAGACCCGGGCGCGCTGTCCAAAGCAGCTGCAGAGCGGCCAGGCC	422
QY	121	ArgLeuGlnAlaAspMetGIuAspValCysGIuArgLeuValGlnTyrArgGIuVal	140
Db	422	CGGCTGGGGCGGCAATGGAGAGAGTGTGGCGCGCTGTGGCAGTACCCGCGGAGAGTGG	484
QY	141	GlnAlaMetLeuGlnGlnSerThrGIuGIuLeuAArgValArgLeuAlaSerHisLeuAArg	160
Db	482	CAGGCCATGGCTGGGCCAGAGACCGAGAGACTGGGGTGGCGCTCCGCGCTCCACACTGGC	544
QY	161	LysLeuCysLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTyr	180
Db	542	AAGCTCGGTAAAGCGGCTTCCTCCGCAATCCGATACCTGCAGAGAAGCGCTGGCAGTAC	604
QY	181	GlnAlaLeuAlaAArgGIuGIuAlaGIuAArgGIuLeuSerAlaLeuArgGIuAArgLeuGIu	200

[illegible]

RESULT 2

Sequence 5, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
TITLE OF INVENTION: Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMM:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:

```

NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 1,65e-102
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4
Matches: 1126
Conservative: 223
Indels: 45
Gaps: 3

US-09-827-854-18 (1-317) x US-08-949-155-5 (1-1126)
QY 1 MetValValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 51 ATGAGGGTCTGTGGTCTTGTGGTACCTCTCGACAGATGCGGACAGAGGAC 110
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 111 GAGCCGGGG-----CCGCCCGGAGAGGTGACGTGTGTGGAGAGAGCCCAAG 158
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
Db 159 TGGCAGGCGACGACGCCCTGGGAGGAGCCCTGGCCCTCTCGGATTAACCTGGCCTGG 218
QY 58 ValGlnThrLeuSerGluGlnAlaGlnGluGlnGluLeuSerSerGlnValThrGlnGlu 77
Db 219 GTGCACTCCCTCTGTACCAAGTGCAGAGAGAGTGTCTGACACCAAGGTCAACCCAGGAA 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
Db 279 CTGACGAGACTGTATAGAGAGAGACATGAGAGGTGAAGGCTACCGCCAGAGAGCTGGAG 338
QY 98 GluGlnLeuThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAla 117
Db 339 GCGCAGCTCGGCGCCCTGTGACCGAGAGAGCGCGCCCTGTCCAAAGAGCTGACGAGCG 398
QY 118 AlaGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
Db 399 GCGCAGGCGCGGTGGGCGCCGACATGAGAGAGGTGCGCAACCGCTGTGTCTTACCCG 458
QY 138 GlyGluValGlnAlaMetLeuGlnLysSerThrGluGlnLeuArgValArgLeuAlaSer 157
Db 459 AGCGAGGTGCACACATGTTGGGCCGACACCGAGAGACTGCGAGCGCCCTGGCTTCC 518
QY 158 HisLeuArgLysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
Db 519 CACCTGCGCAAGCTGCGCAAGCGGCTGCCGCCGACACCGAGACCTGCAGAAAGCCCTG 578
QY 178 AlaValTyrGlnAlaGlnAlaArgGlnGlnAlaArgGlyLeuSerLysAlaLeuArgLys 197
Db 579 GCGGTGTACAGGCGGGGTGGCGGAGGCGCGAGCGAGCTGAGTCCCTCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAla 217
Db 639 CGCCTCGGGCCCTGTGTGAGAGAGGCGCATGTGGCGCGCCACCTGTAGTACAGAGGCC 698
QY 218 GlyGlnProLeuGlnGlnArgAlaGlnAlaTrrpGlyGlnArgLeuArgAlaArgMetGlu 237
Db 699 GGCACACCCCTCGCGGAGCGCGGAAAGCTGTGGCGCAAGACTGCGCGAGCGGTGAG 758
QY 238 GluMetGlySerArgTrpArgAspArgLeuAspGluValLysGlnGlnAlaAlaGluVal 257
Db 759 GAGATGGGAGCGCGGACCGCGACCGCTGTGATGATGATGATGATGATGATGATGATG 818
QY 258 ArgAlaLysLeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnLysAla 277
Db 819 CGGACCAAGGTGAGAGAGAGAGGACCGCATGTGGCTCGAGGCCAGGAGATTCCAGCGC 878
QY 278 ArgLeuLysSerTrpPheGlnProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeu 297
Db 879 CTCCTCAAGGCTGTGAGCTGTGAGGAAAGACATACGCGCGCAGGAGGAGGAGCGGCTG 938

```

```

QY 298 ValGluLysValGlnAlaValAlaGly---ThrSerAlaAlaProValProSerAspAsn 316
Db 939 GTGGAGAGATGATGATGGCGGTGAGACATATAGCTCTCTCCACTTGTGCGCCAGTATAT 998

RESULT 3
US-09-819-964-5
Sequence 5, Application us/09819964
Patent No. 6369294
GENERAL INFORMATION:
Applicant: Pledrahita, Jorge A
Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
Generation of Transgenic Animal Species
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-819-964-5

Alignment Scores:
Pred. No.: 1,65e-102
Score: 1130.00
Percent Similarity: 83.44%
Best Local Similarity: 69.69%
Query Match: 71.11%
DB: 4
Matches: 1126
Conservative: 223
Indels: 45
Gaps: 3

US-09-827-854-18 (1-317) x US-09-819-964-5 (1-1126)
QY 1 MetValValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 51 ATGAGGGTCTGTGGTCTTGTGGTACCTCTCGACAGATGCGGACAGAGGAC 110
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 111 GAGCCGGGG-----CCGCCCGGAGAGGTGACGTGTGTGGAGAGAGCCCAAG 158

```

```

QY 38 TrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGGCAGCCAGCCCTGGAGCAGAGCCCTGGCGCTCTGGGATTAACCTCGCTGG 218
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGAGAGCCCTGCTGACCAAGTGCAGAGAGAGCTGCTGACAGCACCAGAGTACCAGGAA 278
QY 78 LeuArgAlaLeuMetAspGlnThrMetGlyGlnLeuLysAlaTyrLysSerGlnLeuGlu 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCGATAGAGAGAGAGATGAGAGAGTGAAGGCTTACCCGAGAGACTGGAG 338
QY 98 GlnGlnLeuThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGAGAGCTGGGCCCTGGACCCAGAGACAGCGCGGCTGCTCCAGAGAGAGCTGACAGCG 398
QY 118 AlGlnAlaThrLeuGlnValAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGAGAGCCCGCGCTGGCGCGAGACATGAGAGAGCTGCGCAACCGCTTGTGCTCTACCGC 458
QY 138 GlyGlnValGlnAlaMetLeuGlnGlyGlnSerThrGlnGlnLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCAGAGTGCACACATGTTGGGCCACACACAGAGAGCTGGGAGCCGCTGCTTCC 518
QY 158 HisLeuArgLysLeuGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTGGCAGAGCTGCGCAAGCGGCTCTCCGACACCGAGACCTGCGAGAACCGCTTG 578
QY 178 AlaValTyrGlnAlaGlnValAlaArgGlnGlyValGlnGlnGlnLeuSerAlaIleArgGlu 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCCGTGACCAAGCGGGGCTGCGCGAGGCGGCCAGCGCAGCTGAGCCCTCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCCTCGGGCCCTGGTGGAGAGAGAGCGGATTTGGCCCGCAGCTGAGTACAGAGGCGC 698
QY 218 GlyGlnProLeuGlnGlnValArgAlaGlnAlaTyrGlyGlnArgLeuArgMetGlu 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGCGCTGCGCGAGCGCGCGAGCCTGGGCGCAGAACCGCGCGAGGCTGGAG 758
QY 238 GlnMetGlySerArgTyrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGCGAGCCGAGCCCGAGCCGCTGGATGAGATGCTGAGCGAGCTGGAGAGAGTG 818
QY 258 ArgAlaLysLeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaAlaPheGlnAla 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CGCACAAGTGGAGAGAGAGCGAGCGAGTTCGCGCTGCGAGCGCGAGGATTCACAGCC 878
QY 278 ArgLeuLysSerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTyrPheGlnLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCTCAAGGCTGGTTCGAGCTCTGTTGGTGAACATACAGCTGAGTGGCGGCGGCTG 938
QY 298 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGAGAGAGATGACAGTGGCGCTGAGCATTAAGCTCTCCACCTGTGGCGCCAGTGAAT 998

```

```

RESULT 4
US-08-949-155-51
; Sequence 51, Application US/08949155
; Patent No. 6271436
; GENERAL INFORMATION:
; APPLICANT: Pledranta, Jorge A
; APPLICANT: Bazer, Fuller W
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hidler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TAMR:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

```

```

Alignment Scores:
Pred. No.: 2.53e-88 Length: 4267
Score: 994.00 Matches: 217
Percent Similarity: 58.20% Conserved: 42
Best Local Similarity: 48.76% Mismatches: 47
Query Match: 62.56% Indels: 140
DB: 4 Gaps: 5

```

US-09-827-854-18 (1-317) x US-08-949-155-51 (1-4267)

```

QY 1 MetLysValLeuThrPheAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGGCTGTGAACAGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2489
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnProGlnProGlnProGlnProGlnPro 37
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCGCGGGG-----CCGCGCGCGAGAGTGCACGCTGTGTGGAGAGAGCCAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGGCAGCCAGCCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2597
QY 58 ValGlnThrLeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGAGAGCCCTGCTGACCAAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGCACCCAGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCTCTCGGCAACCGTGTGTTGACCTCAGAGCTCAGAGCTCCAGCCGCTCCGCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTTGTCCCAACTCTTGGGGGCTGGGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2836
QY 78 ----- 78
Db 2837 TTGGGGGCAAAACCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTCC 2896
QY 78 ----- 78

```


Qy	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuAlaValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCATGCTCTCGCCAGAGCACCAGAGAGCTGCGGGTGGCCCTCGCTCCACCTGGC	541
Qy	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	542	AAGCGGTAGAGGGCTCTCTCCGGATCCCGATATACCTGGAGAAAGCCCTGGCAGTGTAC	601
Qy	181	GlnAlaGlyAlaArgGlnGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeu	199
Db	602	CAGCGCGGGGCGCCCGAGAGCGCGAGAGCGGGCTCAGCGCCATCCGCGAGCGCCTG	658
RESULT 7			
	US-07-849-389-6		
	Sequence 6, Application US/07849389		
	Patent No. 5525493		
	GENERAL INFORMATION:		
	APPLICANT: HORRES, Erik		
	APPLICANT: UHLEN, Mathias		
	TITLE OF INVENTION: CLONING METHOD AND KIT		
	NUMBER OF SEQUENCES: 7		
	CORRESPONDENCE ADDRESS:		
	ADDRESSEE: Foley & Lardner		
	STREET: 1800 Diagonal Road,		
	City: Alexandria		
	STATE: Virginia		
	COUNTRY: USA		
	ZIP: 22313-0299		
	COMPUTER READABLE FORM:		
	MEDIUM TYPE: Floppy disk		
	COMPUTER: IBM PC compatible		
	OPERATING SYSTEM: PC-DOS/MS-DOS		
	SOFTWARE: PatentIn Release #1.0, Version #1.25		
	CURRENT APPLICATION DATA:		
	APPLICATION NUMBER: US/07/849,389		
	FILING DATE: 19920519		
	CLASSIFICATION: 435		
	ATTORNEY/AGENT INFORMATION:		
	NAME: BENT, Stephen A.		
	REGISTRATION NUMBER: 29,768		
	REFERENCE/DOCKET NUMBER: 16787/168/DFBC		
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE: (703)836-9300		
	TELEFAX: (703)683-4109		
	TELEX: 899149		
	INFORMATION FOR SEQ ID NO: 6:		
	SEQUENCE CHARACTERISTICS:		
	LENGTH: 330 base pairs		
	TYPE: NUCLEIC ACID		
	STRANDEDNESS: single		
	TOPOLOGY: linear		
	MOLECULE TYPE: DNA (genomic)		
	ORIGINAL SOURCE:		
	ORGANISM: HUMAN LIPROTEIN E GENE		
	FEATURE:		
	NAME/KEY: CDS		
	LOCATION: 1..330		
	US-07-849-389-6		
Alignment Scores:			
	Pred. NO.:	1,79e-41	
	Score:	503.00	330
	Percent Similarity:	97.27%	Matches: 106
	Best Local Similarity:	96.36%	Conservative: 1
	Query Match:	31.66%	Mismatches: 3
	DB:	1	Indels: 0
		Gaps:	0
Qy	87	LysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeuThrProValAlaGlnGln	106
Db	1	AAGAGGTGAAGGCTCAATCGAATCGAGACAACTGACCCCGTGGCGGAGAGAG	60

```

Qy      107  |||||Th|||r||g||a||a||r||g||l||e||u||s||e||r||y||s||g||l||e||u||g||l||a||l||a||g||l||n||l||a||r||g||l||e||u||g||l||a||a||s||p||m||e||t|| 126
Db      61  ACGGGGGACGGCGTGTCCAAAGACTCTCAGGCGGCGAGGCCCGCGTGGGGCGGGACATG 120
Qy      127  G|U|A|S|P|A|C|y|S|G|Y|A|G|L|e|V|a|I|n|T|y|A|r|g|I|y|G|l|u|V|a|I|n|a|l|a|e|t|l|e|u|G|I|g|L|n 146
Db      121  GAGGACCGTGGCGGGCCCGCTGTGTGCTACTACCGCGCGCAGGTGCAGGCGCATCTCGGCCAG 180
Qy      147  SerThrglUglUeuArVAlaArgLeuAlaSerHisLeuArglYsUeCysLysArgLeu 166
Db      181  AGCACCGAGAGCTGCGGGTGCCTGCGCTCCACCTCCAGCTGCAGAACTGCTAGACGGGCTC 240
Qy      167  LeuArGAsPaAlaAsPaPleuGlnLysArgLeuAlaValIyTrGlnAlaGlyAlaArgL 186
Db      241  CTCGGCGATGCCGATGTGACCTGCACAAACGGCTGTGCACTGTACCAAGCGGGGGCCCGGAG 300
Qy      187  G|Y|A|L|a|G|l|U|A|r|G|Y|L|e|U|S|e|r|A|l|l|e|A|r|g 196
Db      301  GGGCGGAGCGCGGCGCTCAGCGCATCCGC 330

RESULT 8
US-08-617-256-24
: Sequence 24, Application US/08617256
: Patent No. 6043031
:
: GENERAL INFORMATION:
: APPLICANT: Kvster, Hubert
: TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street, suite 510
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109-1875
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/617,256
: FILING DATE: March 18, 1996
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: March 17, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Arnold, Beth A.
: REGISTRATION NUMBER: 35,430
: REFERENCE/DOCKET NUMBER: SQI-013CP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ. ID NO: 24:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: cDNA
:
: US-08-617-256-24

Alignment Scores:
Pred. No.: 3,87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
DB: Gaps: 0

US-09-827-854-18 (1-317) x US-08-617-256-24 (1-252)

```


MOLECULE TYPE:	CDNA		
US-09-431-613-24			
Alignment Scores:			
Pred. No.:	3.87e-28	Length:	252
Score:	366.00	Matches:	81
Percent Similarity:	96.43%	Conservative:	0
Best Local Similarity:	96.43%	Mismatches:	3
Query Match:	23.03%	Indels:	1
DB:	4	Gaps:	0
US-09-827-854-18 (1-317) x US-09-431-613-24 (1-252)			
QY	109	AAaargleuSerlySGluLeuGlnAlaAlaGlnAlaargleuGlyAlaAspMetGluAsp	128
Db	2	GCAGCGCTGTCCAAGAGCATCGACGCGCGAGCCGCGGCTGGCGCGGACATGAGAGAC	61
QY	129	ValcysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr	148
Db	62	GTGTGCGC-CGCTGCTGTCACTACCGCGCGAGGTGAGGCAATGCTGGCGCAAGCACC	120
QY	149	GluGluLeuArgValArgLeuAlaSerHisLeuArgIysLeuGlySylsArgLeuArg	168
Db	121	GAGGAGTGCGGGGTTGGCGCTGCTCCACCTGCGCAAGCTGGCTAAGCGGCTCTCGC	180
QY	169	AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla	188
Db	181	GATGCCGATGACCTGCAGAAATCCCTCGACGTATCACAGCGCGGCGCCGAGGGCGCC	240
QY	189	GluArgGlyLeu	192
Db	241	GAGCGCGGCTTC	252
RESULT 11			
US-09-504-245-24			
Sequence 24, Application US/09504245			
Patent No. 6221605			
GENERAL INFORMATION:			
APPLICANT: K Ster, Hubert			
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry			
NUMBER OF SEQUENCES: 33			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Heller Ehrman White & McAlliff LLP			
STREET: 4250 Executive Square, 7th Floor			
CITY: La Jolla			
STATE: CA			
COUNTRY: USA			
ZIP: 92037-9103			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette			
COMPUTER: IBM Compatible			
OPERATING SYSTEM: DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/504,245			
FILING DATE:			
CLASSIFICATION:			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/617,256			
FILING DATE: 18-MAR-1996			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/406,199			
FILING DATE: 03-MAR-1995			
ATTORNEY/AGENT INFORMATION:			
NAME: Seidman, Stephanie L			
REGISTRATION NUMBER: 33,779			
REFERENCE/DOCKET NUMBER: 24736-2002J			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 858-450-8400			
TELEFAX: 858-587-5360			
INFORMATION FOR SEQ ID NO: 24:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 252 base pairs			

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-504-245-24

Alignment Scores:
Pred. No.:          3.87e+28           Length:        252
Score:              366.00             Matches:         81
Percent Similarity:   96.43%            Conservative:    0
Best Local Similarity: 96.43%            Mismatches:      3
Query Match:         23.03%             Indels:          1
DB:                  4                   Gaps:            0

US-09-827-854-18 (1-317) x US-09-504-245-24 (1-252)
Oy       109 AlaArgLeuSerLySGluL euGlnAlA lalAGl nlaArgrleugLiYALaaSPmeGLuaSp     128
||| |||| |
Db       2 GCACGGCGCTGTCCAAGAGACTCAGAGCCGCCGACC GCCGCTGGCGCGGCATGGAAGAAC     61
||| |||| |
Oy       129 ValCSglYA rgleuValGI ntTYr ArgLY glYuVa lGInal meLteugLI ynS erThr     148
||| |||| |
Db       62 GTTGTCGC -CCGCTGGTCACTAC TCCGCGG CAGA GGSTG CAGGCC ATG CTGG CCACA GA CACC     120
||| |||| |
Oy       149 GlUgtLU eUA rV alAr gLE uAl A SerHis tEu ar gLy sE ucY lySA rGLE uLa rg     168
||| |||| |
Db       121 GAGGAGCGCTGCGGGGTGCGCTTGCCTCCACCTGCGCAAGCTGCTAA GCGGCTCTCCG     180
||| |||| |
Oy       169 AsPalaasPaSPleUGln LysAr gLE uAl aVal Tyrg InaI aglyAL aa rGr Lu gLY AL a     188
||| |||| |
Db       181 GATGCCCATGACCTTCGACAAAGT CCTTGCCTG CAGTA TCAC CA GCG CGG GCGCCCG AGGGCC     240
||| |||| |
Oy       189 GlUAR gLY leu 192
||| |||| |
Db       241 GAGCGCGGCGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent NO. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McLaughlin
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 3.87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-682-24 (1-252)

QY 109 AAlarqleuserlysgluLeuGlnAlaGlnAlaargleuGlyAlaaspmetGluasp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCTGGCGGACATGAGAGAC 61

QY 129 ValcysGlyArGleuValGlnTyrArGlyGlyValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGGTACCTACCGGCGGAGGTGAGGCGCATGTCTGGCAGAGAC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArGlyLeuGlySerArGleuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCTCTGCTCCACCTCGCAAGCTGCTAAGCGGCTCTCTCGC 180

QY 169 AspaAlaaspLeuGlnlyArGleuAlaValTyrGlnAlaGlyAlaargGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGGCAGTGTACAGGCGCGGCGCGGAGGCGCC 240

QY 189 GluArGlyLeu 192
Db 241 GAGCGCGGCTCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 3.87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-679-24 (1-252)

QY 109 AAlarqleuserlysgluLeuGlnAlaGlnAlaargleuGlyAlaaspmetGluasp 128
Db 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGCGCTGGCGGACATGAGAGAC 61

QY 129 ValcysGlyArGleuValGlnTyrArGlyGlyValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGGTACCTACCGGCGGAGGTGAGGCGCATGTCTGGCAGAGAC 120

QY 149 GluGluLeuArGValArGleuAlaSerHisLeuArGlyLeuGlySerArGleuLeuArg 168
Db 121 GAGGAGCTGGGGGTGGCTCTGCTCCACCTCGCAAGCTGCTAAGCGGCTCTCTCGC 180

QY 169 AspaAlaaspLeuGlnlyArGleuAlaValTyrGlnAlaGlyAlaargGluGlyAla 188
Db 181 GATCCCATGACCTGCAGAGAGTCCCTGGCAGTGTACAGGCGCGGCGCGGAGGCGCC 240

QY 189 GluArGlyLeu 192
Db 241 GAGCGCGGCTCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 3,87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-18 (1-317) x US-09-397-766-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGGCGGCGGCTGGGCGCGGAGCATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCG-CGGCTGTGTGAGTACCGCGGCGAGGTGCAGGCATGCTCGGCCAGAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuGlnCysLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCGCTCCACCTGCCAGAGCTGGTAAAGGGGCTCCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGGCAGGTGTACAGGCGGGGCCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCGCTC 252

RESULT 15
US-09-287-681-24

; Sequence 24, Application US/09287681
; Patent No. 6277573
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 3,87e-28 Length: 252
Score: 366.00 Matches: 81
Percent Similarity: 96.43% Conservative: 0
Best Local Similarity: 96.43% Mismatches: 3
Query Match: 23.03% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-18 (1-317) x US-09-287-681-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGGCGGCGGCTGGGCGCGGAGCATGGAGGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCG-CGGCTGTGTGAGTACCGCGGCGAGGTGCAGGCATGCTCGGCCAGAGACACC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuGlnCysLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCGCTCCACCTGCCAGAGCTGGTAAAGGGGCTCCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAAATCCCTGGCAGGTGTACAGGCGGGGCCGCGAGGGCGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCGCTC 252

Search completed: March 14, 2003, 20:18:51
Job time : 39.7343 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVLMALVTLFLAGCQAKV.....VEKYQAAVGTSAAPVSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09827854/rumat_11032003.101612.27578/app_query.fasta.1.3576
-DB=Published_Applications_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=dits -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854@cgn2.1.1.298@rumat_11032003.101612.27578
-NCPU=6 -ICPU=3 -NO_XLIPX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGSLOG
-DEV_TIMEROUT=120 -WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

Database : Published Applications_NA:*

- 1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2.6/ptodata/1/pubpna/PCY_NEW_PUB.seq:*
- 3: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2.6/ptodata/1/pubpna/PCYUS_PUBCOMB.seq:*
- 7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1589	100.0	1156	10	US-09-827-854-11
2	1577	99.2	1156	9	US-09-870-759-129
3	1577	99.2	1156	9	US-09-802-640-17
4	1577	99.2	1156	10	US-09-827-854-8

5	1577	99.2	1291	12	US-10-044-090-454	Sequence 454, App
6	1573	99.0	1156	10	US-09-827-854-12	Sequence 12, App
7	1569	98.7	1156	10	US-09-827-854-9	Sequence 9, Appl
8	1568	98.7	1157	10	US-09-954-456-760	Sequence 760, App
9	1568	98.7	1157	10	US-09-880-107-2244	Sequence 2244, App
10	1565	98.5	1156	10	US-09-827-854-7	Sequence 7, Appl
11	1562	98.3	1156	10	US-09-827-854-10	Sequence 10, Appl
12	1368.5	86.1	41907	10	US-09-967-013-5	Sequence 5, Appl
13	877.5	55.2	786	10	US-09-925-302-133	Sequence 133, App
14	625	39.3	478	10	US-09-964-824-374	Sequence 374, App
15	625	39.3	478	10	US-09-880-107-2491	Sequence 2491, App
16	458.5	28.9	449	10	US-09-960-352-4726	Sequence 4726, App
17	455	28.6	356	10	US-09-960-352-5420	Sequence 5420, App
18	439.5	27.7	442	10	US-09-960-352-9395	Sequence 9395, App
19	435	27.4	414	10	US-09-960-352-1237	Sequence 4237, App
20	428.5	27.0	416	10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.6	423	10	US-09-960-352-14047	Sequence 14047, A
22	405.5	25.5	409	10	US-09-960-352-5148	Sequence 5148, App
23	401.5	25.3	416	10	US-09-960-352-9797	Sequence 9797, App
24	393	24.7	253	10	US-09-179-5368-130	Sequence 130, App
25	387.5	24.4	425	10	US-09-960-352-3497	Sequence 3497, App
26	380.5	23.9	390	10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391	10	US-09-960-352-1278	Sequence 1278, App
28	379.5	23.9	392	10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401	10	US-09-960-352-1187	Sequence 4187, App
30	375.5	23.6	413	10	US-09-960-352-8042	Sequence 8042, App
31	369.5	23.3	404	10	US-09-960-352-3235	Sequence 3235, App
32	369.5	23.3	405	10	US-09-960-352-14063	Sequence 14063, A
33	366	23.0	252	10	US-09-796-416-24	Sequence 24, Appl
34	366	23.0	252	10	US-09-879-341-24	Sequence 24, Appl
35	351.5	22.1	337	10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	353	10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378	10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	386	10	US-09-960-352-10361	Sequence 10361, A
39	348.5	21.9	388	10	US-09-960-352-11966	Sequence 11966, A
40	344.5	21.7	396	10	US-09-960-352-10457	Sequence 10457, App
41	344.5	21.7	377	10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.3	432	10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332	10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.2	376	10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.1	373	10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
GENERAL INFORMATION:
; APPLICANT: Zannits, Vassilis
; TITLE OF INVENTION: Kypreos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-11

Alignment Scores:
Pred. No.: 4.75e-138 Length: 1156
Score: 1589.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0


```
Db 726 CAGGCCGGGGCCCGAGGGCCCGAGCCGCGCTCAGCGCCATCCGCGAGCCGCTGGGG 785
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 786 CCCCTGTGTGAACAGGGCCGCGTGGGGCCGACATGTGTGGCTCCCTGGCGGGCCAGCG 845
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 846 CTACAGAGAGCGGGCCAGGCGCTGGGGCGAGGGCTGCGGGCGGGGATGAGAGAGATGGGC 905
Qy 241 SerArgThrArgAspArgLeuAspGluValGlyGluGlnValAlaGluValArgAlaLys 260
Db 906 AGCCGACCCCGGACCGCTGGACGAGTGAAAGACAGAGAGAGAGAGAGAGAGAGAGAG 965
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 966 CTGGAGAGAGAGCCCGACGAGATACGCTGACAGCCGAGGAGCCCTCCAGGCGCGCTCAAG 1025
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
Db 1026 AGCTGTTCGAGCCCTGTGGTGAAGACATGACAGCCGACAGTGGGGCGGGCTGTGGAGAG 1085
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 1086 GTGCAGGCTGCCTGGGACACACGCGCCGCTGTGCCAGCAGCAATCAG 1136

RESULT 6
US-09-827-854-12
; Sequence 12, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-12

Alignment Scores:
Pred. No.: 1,42e-136
Score: 1573.00
Percent Similarity: 99.68%
Best Local Similarity: 99.37%
Query Match: 98.99%
DB: 10
Matches: 1156
Conservative: 1
Mismatches: 1
Indels: 0
Gaps: 0

US-09-827-854-18 (1-317) x US-09-827-854-12 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATCAAGGTTCTGTGGCTGCTGTGCTGTGCATCTCTCGGAGAGATGCCAGGCAAGTGG 120
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTCGCCAGAGAGACCAAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 181 GGCACAGCTGGGACATGCGACTGGTCTTTGGATTAACCTGCGTGGGGGAGACA 240
Qy 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
```

```
Db 241 CTGTCTGACAGAGTGCAGAGAGAGTGTCTGACGTCCAGATCCAGAGAACTGAGGGC 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGAGAGACATGAAAGAGTGTGAAGGCTTAAATAATCGGAATGAGAACTATG 360
Qy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGGAGAGAGACGCGGGACAGGCTCTCAAGAGAGCTGCAGCGGGCGAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 421 CGCGTGGCGCGGACATGAGAGAGAGTGTGGCGCCGCTGGTGTGACATCCGCGGAGAGTGG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCAATGCTCTGCGCAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGCTGACAGCGCTCTCCGCGATGCGGATGACCTGCAGAAAGCCCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGGCGCGAGGGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 661 CCCCTGTGTGAACAGGGCCGCGTGGGGCCGACATGTGGGCTCCCTGGCGGGCCAGCGG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGGCGCTGCGCGGGATGAGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValGlyGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGACCCGACCGCTGCGAGAGTGTAAGACAGAGTGGCGGAGGCGCGCCAGAG 840
Qy 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGACAGGCCCGACGATACCTGTGACAGCCGAGGCTTCCAGGCGCGCTTCAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
Db 901 AGCTGTTCGAGCCCTGTGGTGAAGACATGACAGCCGCAATGGGCGCGGCTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCCTGGGACACACGCGCCGCTGTGCCAGCAGCAATCAG 1011

RESULT 7
US-09-827-854-9
; Sequence 9, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
```



```

|||||
Db 362 ACCCGGTACGAGGAGAGCGCGGCGACGCTCTCCAGAGAGCTGACAGCGCGAGGCC 421
QY 121 ArgLeuGlyValAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGCGTGGCGCGGACATGGAGGACGTGTGGCGCGCTGGTGTGATACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGGCGCAGAGCACCGAGGAGCTCGGGTGGCTCCGCTCCACCTGCGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTCGTAAAGCGGCTCTCCCGCGATGCCGATGACCTGACGAAGCGCTGCGAGGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCGCGCAGGGCGCGCAGCGGCTGACGCGCCATCCGCGAGCGCGTGGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGGTGAACAGGGCGCGTGGCGCGCCACTGTGGGCTCCCTGGCGCGCGACCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGGATGGAGAGTGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGluValAlaGluValArgAlaLys 260
Db 782 AGTCGGACCCCGCAGCGCTCGGACGAGGTAAAGAGCAGGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGAGCGCCAGAGATAGCTGTGCAAGCGGAGGCTTCCAGGCGCGCGCTCAAG 901
QY 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
Db 902 AGCTGGTTCGAGCCCTGTGGTGAAGACATGACAGCGCAGGCGCGGCTGTGGAGAAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaIleAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCCGTGGGACACAGCGCGCGCTGTGCCAGGACAAATCAC 1012

RESULT 9
US-09-880-107-2244
; Sequence 2244, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2244
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

```

```

|||||
Db 362 ACCCGGTACGAGGAGAGCGCGGCGACGCTCTCCAGAGAGCTGACAGCGCGAGGCC 421
QY 121 ArgLeuGlyValAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGCGTGGCGCGGACATGGAGGACGTGTGGCGCGCTGGTGTGATACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGGCGCAGAGCACCGAGGAGCTCGGGTGGCTCCGCTCCACCTGCGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTCGTAAAGCGGCTCTCCCGCGATGCCGATGACCTGACGAAGCGCTGCGAGGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 602 CAGGCGGGGGCGCGCAGGGCGCGCAGCGGCTGACGCGCCATCCGCGAGCGCGTGGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGGTGAACAGGGCGCGTGGCGCGCCACTGTGGGCTCCCTGGCGCGCGACCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGGATGGAGAGTGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGluValAlaGluValArgAlaLys 260
Db 782 AGTCGGACCCCGCAGCGCTCGGACGAGGTAAAGAGCAGGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGAGCGCCAGAGATAGCTGTGCAAGCGGAGGCTTCCAGGCGCGCGCTCAAG 901
QY 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
Db 902 AGCTGGTTCGAGCCCTGTGGTGAAGACATGACAGCGCAGGCGCGGCTGTGGAGAAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaIleAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCCGTGGGACACAGCGCGCGCTGTGCCAGGACAAATCAC 1012

RESULT 10
US-09-827-854-7

```

Percent Similarity: 99.05%
 Best Local Similarity: 99.05%
 Query Match: 98.68%
 DB: 10
 Gaps: 0

US-09-827-854-18 (1-317) x US-09-880-107-2244 (1-1157)

```

QY 1 MetLysValLeuTrrpAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGGCTGCGTGTGTCACATTCTTGGCAGAGATGCCAGGCGCAAGGTG 121
QY 21 GlnGlnAlaValGlnThrGlnProGlnLeuArgGlnGlnThrGlnSer 40
Db 122 GACCAAGCGGTGAGACAGAGCCGAGCCGAGACTCGCCAGCAGACAGCAGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrrpGluLeuAlaLeuGlyArgPheTrrpAspTrrpLeuArgTrrpValGlnThr 60
Db 182 GGCACGCGCTGGGACATGGCATGGGTGCTCTTGGGATTTACCTGCGCTGGTGGCAGACA 241
QY 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGACAGAGTGCAGAGAGAGCTGTCTACGCTCCCAAGTCCCAAGAGACTGAGAGGCC 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrpLysSerGluLeuGlnGlnLeu 100
Db 302 CTGATGAGAGACACATGAAGGAGTTGAAGGCTTCAAAATCCGAACCTGAGAGCAAACTG 361
QY 101 ThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIleAlaGlnAla 120
Db 362 ACCCGGTAGCGGAGAGAGCGCGCGCGCTGTCCAGAGCTGAGACAGCGCGCAGGCC 421
QY 121 ArgLeuGlyValAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCGCTGTGTGAGTACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTCGGCGCAGAGCACCGAGAGCTCGGGTGGCGCTCCCTGCCCTCCACTGCGC 541
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AACCTCGTAAAGCGGCTCTCCCGCGATGCCGATGACCTGACGAAGGCGCTGGAGGTGAC 601
QY 181 GlnAlaGlyAlaArgGlnGlnIleArgGlnAlaGlnAlaIleArgGlnAlaArgLeuLys 200
Db 602 CAGCGCGGGGGCGCGCAGGGCGCGCAGCGGCTTCAAGCGCGCGCGCGCTGCGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
Db 662 CCCCTGGTGAACAGAGCGCGCGCGGCGCGCACTGTGGGCTCCCTGGCGCGCGACCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaTrrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 722 CTACAGAGAGCGGCGCCAGGCTGGGGCGAGCGGCTGCGCGGGTGAAGAGATGGGTC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGluValAlaGlnAlaArgAlaLys 260
Db 782 AGTCGGACCCCGCAGCGCTCGGACGAGGTGAAGAGCAGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGAGCGCCAGAGATAGCTGTGCAAGCGGAGGCTTCCAGGCGCGCGCTCAAG 901
QY 281 SerTrrpHeGluProLeuValGluAspMetGlnArgGlnTrrpAlaGlyLeuValGluLys 300
Db 902 AGCTGGTTCGAGCCCTGTGGTGAAGACATGACAGCGCAGGCGCGGCTGTGGAGAAAG 961
QY 301 ValGlnAlaAlaValGlyThrSerAlaIleAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCCGTGGGACACAGCGCGCGCTGTGCCAGGACAAATCAC 1012

RESULT 10
US-09-827-854-7

```

Alignment Scores: 4.11e-136
 Pred. No.: 1568.00
 Length: 1157
 Matches: 314

```

; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacs, Kyriacos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7

```

Alignment Scores:

```

Pred. No.: 7,76e-136      Length: 1156
Score: 1565.00             Matches: 315
Percent Similarity: 99.37%  Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 98.49%        Indels: 0
DB: 10                     Gaps: 0

```

US-09-827-854-18 (1-317) x US-09-827-854-7 (1-1156)

```

QY 1 MetLysValLeuTPRALaLaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||
DB 61 ATGAGGTTCTGTGGCTGGCTGCTGTCTGTCACTCTGGAGAGATCCAGGCCAAGTG 120
    |||
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGlnSer 40
    |||
DB 121 GAGCAACGGGTGGAGACAGCCGAGCCGAGCTGCCAGCCAGACCGAGTGCCAGAGC 180
    |||
QY 41 GlyGlnArgTPRGLuLeuAlaLeuGlyArgPheTPRSPryLeuArgTPRValGlnThr 60
    |||
DB 181 GGCACAGCGCTGGAGACTGGACTGGCTGCTTTGGGATTTACCTGCGCTGGGTGCAGACA 240
    |||
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||
DB 241 CTGCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGCAAGAACTAGGGCG 300
    |||
QY 81 LeuMetLysPgluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
    |||
DB 301 CTGATGGACAGACCATGAAGAGATTGAAGCCCTACAAATCGGAATCGAGGAACAACGTG 360
    |||
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
    |||
DB 361 ACCCGGTGGCGGAGAGACCGGGGACGGCTGTCCAAAGAGCTGCAGGGCGGCCAGGCC 420
    |||
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||
DB 421 CGGCTGGCGCGGACATGAGAGAGCTGCGCGCGCTGAGTGCAGTACCGGGGAGAGTG 480
    |||
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||
DB 481 CAGGCACTGCTGGCGCCAGACACCGAGAGCTGTGGGTGGCTCCCTCCACTCGCGC 540
    |||
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||
DB 541 AAGCTGGCTTAAGCGGCTCTCCCGGATGCCGATGACTCTGGAAGGCCCTGGAGCTATC 600
    |||
QY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||
DB 601 CAGGCGGGGCGCGGAGGGCGCGGAGCGGGCTCAGGCCCATCCGCGAGGCGCTGGGG 660
    |||
QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
    |||

```

```

DB 661 CCCCTGGTGAACAGAGCGCGCTCGGCGCCCACTGTGGCTCTCCGTGGCGGCACCG 720
    |||
QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluThrArgLeuArgAlaArgMetGluMetGly 240
    |||
DB 721 CTACAGAGCGGGCCAGGCTGGGGGAGCGGCTGGCGCGCGGAGTGGAGAGATGGGC 780
    |||
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||
DB 781 AGCCGGAACCGCGACCCGCTGTGAGAGTGAAGAGACAGTGGCGGAGAGTGGCGGCAAG 840
    |||
QY 261 LeuGluGluGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAlaArgLys 280
    |||
DB 841 CTGAGAGAGCAGAGCCAGCAGATAGCGCTGAGGCCAGAGGCTTCCAGGCGCGCTCAG 900
    |||
QY 281 SerTPRTPRLeuProLeuValGluAspMetGlnArgGlnTPRAlaGlyLeuValGlyLys 300
    |||
DB 901 AGCTGTTCCAGACCCCTGTGGTGAAGACATGACAGCGCCAGTGGCGGCTGTGGAGAG 960
    |||
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||
DB 961 GTGCAGGCTGCGCTGGGACACAGCGCGCCCTGTGCCACAGCCACATTCAC 1011
    |||

```

RESULT 11

```

US-09-827-854-10
; Sequence 10, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacs, Kyriacos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10

```

Alignment Scores:

```

Pred. No.: 1,47e-135      Length: 1156
Score: 1562.00             Matches: 314
Percent Similarity: 99.05%  Conservative: 0
Best Local Similarity: 99.05% Mismatches: 3
Query Match: 98.30%        Indels: 0
DB: 10                     Gaps: 0

```

US-09-827-854-18 (1-317) x US-09-827-854-10 (1-1156)

```

QY 1 MetLysValLeuTPRALaLaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||
DB 61 ATGAGGTTCTGTGGCTGGCTGCTGTCTGTCACTCTGGAGAGATCCAGGCCAAGTG 120
    |||
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTPGlnSer 40
    |||
DB 121 GAGCAACGGGTGGAGACAGCCGAGCCGAGCTGCCAGACAGACCGAGTGCAGAGC 180
    |||
QY 41 GlyGlnArgTPRGLuLeuAlaLeuGlyArgPheTPRSPryLeuArgTPRValGlnThr 60
    |||
DB 181 GGCACAGCGCTGGAGACTGGACTGGCTTGGGATTTACCTGCGCTGGGTGCAGACA 240
    |||
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||
DB 241 CTGCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGCAAGAACTAGGGCG 300
    |||

```

QY 81 LeuMetaSpclurthrmetylsGluLeuLysAlaTyrLysSerGluLeuGluGluGlu 100
Db CTGATGGACGACCATGAAAGGAGTGAAGGCTTACAAATCGGAACCTGAGGAACAATG 360
QY 101 ThrProValAlaGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db ACCCGGTGGCGGAGAGAGCGCGGCAAGGCTGTCCAAAGAGGTGAGCGGCGGCGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db CGGCTGGCGGCGACATGAGAGACGTGTGGCGGCGCTGGTGCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db CAGGCACTGCTGCACAGACACCGAGAGCTCGGGTGGCGCTCCCTCCACCTGCGG 540
QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db AACCTCGTAAGGCGCTCTCCCGCATGCCGATGCCAGTCCAGAAAGTCCCTGCGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluGluGluArgGlyLeuSerAlaTyrLeuArgGluArgLeuGly 200
Db CAGCGCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db CCCCTGTGGAAACAGGCGCGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db CTTCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db AGCGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 261 LeuGluGluGlnAlaGlnGlnAlaTyrLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db CTGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
QY 281 SerTyrPheGluProLeuValGluAspMetGlnArgGlnTyrAlaGlyLeuValGluLys 300
Db AGCTGTGTCGAGCCCTGTGTGGAGACATGACAGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db GTCAGAGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1011

RESULT 12

US-09-967-013-5
; Sequence 5, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013
; PRIOR FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/206,613
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-013-5

Alignment Scores:

Pred. No.: 7,98e-116
Score: 1368.50
Percent Similarity: 60.00%

Length: 41907
Matches: 300
Conservative: 0

Best Local Similarity: 60.00%
Query Match: 86.12%
DB: 10
US-09-827-854-18 (1-317) x US-09-967-013-5 (1-41907)
Mismatches: 6
Indels: 194
Gaps: 1

QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGlu 31
Db TTCCACACAGAGATGCCAGRCACAGGTGGAGCAAGCGGTGGAGACAGCGGAGCGCGGAG 20375
QY 32 LeuArgGlnGlnThrLutyrGlnSerGlyGlnArgTyrPheGluLeuAlaLeuGlyArgPhe 51
Db CTGCGCCACAGACGACGAGGTGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20435
QY 52 TyrAspTyrLeuArgTyrPheValGlnThrLeuSerGluGlnValGlnGluLeuLeuSer 71
Db TGGGATTACCTGCGCTGGGTGGAGACACTGTCTGACAGGTGACAGAGGAGGAGTGTACG 20495
QY 72 SerGlnValThrGlnGluLeu 78
Db TCCACAGGTACCCAGGACGAGTGTAGTGTCCCATTCCTGAGCCCTTGAACCCCTCTGTG 20555
QY 78 78
Db GGGCGCTATACCTCCCGAGGTCCAGGTTTCATTTCTGCCCCCTGTGCTTGGGCGG 20615
QY 78 78
Db CTTGGGTCTGTGCTGTGTCTAGCTTCCCTCTCCATTTCTGACTCCTGCTTAACTCTTC 20675
QY 78 78
Db TGGGATTCTCTCTCTCAGGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20735
QY 78 78
Db CGTCTGCTCTGTCT 20795
QY 78 78
Db CTCACCTGTGTGCCAGGCTGTCTTGAACCTTGTGGGCTCAGGATCTCTCTCTCTCTCT 20855
QY 78 78
Db CCTCCCAAGAGTGTGGGATTAAGAGCATAGCACCTTGCCCGGCTCTAGTCTCTCTCT 20915
QY 78 78
Db TCGTCTGTGCTCTGTGCTCTGTGATCTGTCTGTGATCTGTCTGTCTCTCTCTCTCTCT 20975
QY 78 78
Db GCCTCTGCCCCGTTCT 21035
QY 78 78
Db CCCCATCCAGGCT 21095
QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 98
Db CAGGCGGCTGATGAGACGACCATGAAAGATGTAAGGCTTCAAAATCGGAACCTGAGAGA 21155
QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
Db ACACACTGACCCCGGTGGCGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 21215
QY 118 aglnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGln 138
Db GCAAGCGCGGCTGGGCGGCGGACATGAGAGACGTGTGGCGGCGGCGGCGGCGGCGGCGG 21275
QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHis 158
Db CGAGGTGCGAGGCATGCTGCGGCGGAGACCGAGAGCTGTGGGCGGCGGCGGCGGCGGCG 21335


```

Db 396 ATGAGGCTTCGTGGCGCTGGCTGTGCATTCCTGGCAGAGCCAGCCAAAGTG 317
QY 21 GUGLAlaValGluThrGluProGluLeuArgGluThrGluTrpGlnSer 40
Db 336 GACCAAGCGGTGGAGACACAGCGGAGCCNCCAGCTCGCGCAGAGACCAAGTGGCAGAC 277
QY 41 GUGLArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACACGCTGGGAACTGGCAGCTGGCTTTGGGATTACTGTGGCGTGGGTGCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGTCTGACACAGGTGCAGAGAGGAGCTGCTGCCAGCTCCAGCTCACCAGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
Db 156 GCGATGAGACGAGACCATGAAGAGTTGAAGCCTTCAAAATCGAACTGAGAGAACACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCGGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTCAGCGCGCGGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGCGACATGAGAGACGTG 9

```

RESULT 15

```

US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scheff, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank accession No. US20020142981A1 N33009
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

```

Alignment Scores:

```

Pred. No.: 1.27e-49 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 39.33% Indels: 1
DB: 10 Gaps: 0

```

US-09-827-854-18 (1-317) x US-09-880-107-2491 (1-478)

```

QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAGGCTTCGTGGCGCTGGCTGTGCATTCCTGGCAGAGCCAGCCAAAGTG 337
QY 21 GUGLAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 336 GACCAAGCGGTGGAGACACAGCGGAGCCNCCAGCTCGCGCAGAGACCAAGTGGCAGAC 277

```

```

QY 41 GUGLArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACACGCTGGGAACTGGCAGCTGGCTTTGGGATTACTGTGGCGTGGGTGCAGACA 217
QY 61 LeuSerGluGlnValGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGTCTGACACAGGTGCAGAGAGGAGCTGCTGCCAGCTCCAGCTCACCAGAACTGAGGGC 157
QY 80 AleuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLe 100
Db 156 GCGATGAGACGAGACCATGAAGAGTTGAAGCCTTCAAAATCGAACTGAGAGAACACT 97
QY 100 uThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCGGGTGGCGGAGAGACGCGGCGACGCTGTCCAAAGAGCTCAGCGCGCGGAGN 37
QY 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGCGACATGAGAGACGTG 9

```

Search completed: March 15, 2003, 03:04:53
.. Job time : 81.8753 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-18

Perfect score: 1589

Sequence: 1 MKVIMALLVTFIACGQAKV.....VEKVOAAVTSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09827854/runat_11032003_101610_27495/app.query.fasta_1.3576
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.eccn.1.1.882c@runat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEIOERR -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1282.5	80.7	1027	13	B1670367	B1670367 603292738
2	1246	78.4	842	12	BG763371	BG763371 602735433
3	1240	78.0	933	13	B1668318	B1668318 603295681
4	1212	76.3	938	12	BG761746	BG761746 602717942
5	1209.5	76.1	942	13	B1600906	B1600906 603249241
6	1195	75.2	800	13	BM042094	BM042094 603615713
7	1194.5	75.2	922	12	BG472299	BG472299 602513830
8	1191.5	75.0	922	13	B1597743	B1597743 603248609
9	1182.5	74.4	817	12	BG774871	BG774871 602649975
10	1181	74.3	811	13	B1600563	B1600563 603244936
11	1167.5	73.5	845	12	BG829472	BG829472 602763768
12	1144	72.0	790	12	BG707147	BG707147 602670283
13	1141	71.8	919	13	B1551475	B1551475 603194314
14	1131	71.2	757	13	BM042228	BM042228 603616186
15	1121.5	70.6	907	12	BG706129	BG706129 602669093
16	1120	70.5	706	14	BM728696	BM728696 UI-E-EJO-
17	1120	70.5	741	12	BG762924	BG762924 602735153
18	1118	70.4	1110	11	AK010261	AK010261 Mus. muscu
19	1116	70.2	804	12	BG702752	BG702752 602684616
20	1115	70.2	834	13	BM042676	BM042676 603616028
21	1112.5	70.0	808	13	B1668329	B1668329 603295692
22	1109	69.8	855	13	B1161362	B1161362 602865769
23	1106	69.6	803	13	B1670350	B1670350 603292721
24	1103	69.4	782	12	BG716776	BG716776 602678182
25	1103	69.4	812	13	B1601551	B1601551 603250760
26	1103	69.4	965	14	BG677266	BG677266 AGENCOURT
27	1100	69.2	757	13	B1603658	B1603658 603251982
28	1097.5	69.1	797	12	BG715366	BG715366 602675503
29	1097	69.0	846	13	B1159757	B1159757 6026863509
30	1096	69.0	796	13	BM042153	BM042153 603615790
31	1091	68.7	757	12	BG707750	BG707750 602671277
32	1091	68.7	794	13	B1601279	B1601279 603245010
33	1091	68.7	954	12	BF667543	BF667543 602287404
34	1089	68.5	790	13	B1551066	B1551066 603196425
35	1088	68.5	790	13	B1551811	B1551811 603192716
36	1084.5	68.3	798	12	BG708414	BG708414 602672374
37	1083.5	68.2	891	13	B1549292	B1549292 603189930
38	1077.5	67.8	802	13	B1458355	B1458355 603198705
39	1076	67.7	1100	14	BM914382	BM914382 AGENCOURT
40	1075.5	67.7	914	13	B1603523	B1603523 603244424
41	1075	67.7	748	13	B1553085	B1553085 603196058
42	1074.5	67.6	930	12	BF667857	BF667857 602287176
43	1070.5	67.4	812	12	BG769968	BG769968 602739854
44	1070	67.3	688	13	B1544886	B1544886 603242328
45	1069.5	67.3	803	12	BG709360	BG709360 602673385

ALIGNMENTS

RESULT 1
B1670367
LOCUS 1027 bp mRNA EST 12-SEP-2001
DEFINITION 603292738F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5112024 5',
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS mRNA sequence.
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA library Preparation: Michael J. Brownstein (NHGRI), Shirohaki Toshiyuki and Piero Carninci (RIKEN)
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LRAM11790 row: 1 column: 09
High quality sequence: 845.

Location/Qualifiers
1. .1027

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_1lp="NH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptpr (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTGTTTTGTTTT-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NIHGR, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT
ORIGIN
194 a
308 c 400 g 125 t

```

Pred. No.:	2,156-144	1027
Score:	1282.50	291
Percent Similarity:	90.55%	6
Best Local Similarity:	88.72%	20
Query Match:	80.71%	12
DB:	13	Gaps:
US-09-827-854-18 (1-317) x B1670367 (1-1027)		

QY	1	MethylValLeuThrAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	26	ATTAAGAGTTCTGGGGCTGGCTGTGGTGCATTTCTGGCGAGATGCCAGGCCAAGGTG	85
QY	21	GlnGlnAlaValAlaGlnThrGlnProGlnProLeuLeuArgGlnGlnThrGlnTyrPgiInser	40
Db	86	GACCAAGCGGTGGAGACAGAGCCGAGGCCAGCTGGGCCAGCAGACCGAGTGGCAGAAC	14
QY	41	GlyGlnArgTyrPgiLeuLeuAlaLeuGlnYArgPheThrPaspTyrLeuArgTyrPValGlnThr	60
Db	146	GGCCACGCGTGGAGAACTGGCACTGAGTGTGCTTTTGGGATTACCTGGCGTGGTGAGACA	20
QY	61	LeuSerGlnGlnValGlnGlnGlnLeuLeuLeuSerSerGlnValThrGlnLeuLeuArgAla	80
Db	206	CTGTCTGACACAGCTGCAGAGAGAGTGTCTCAGCTCCAGATGCACCCAGAACAGTGAAGGGCG	26
QY	81	LeuMetAspGlnThrMetLysGlnLeuLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu	10
Db	266	CTGATGACGACACACCATGACGAGTTGAAGGCCCTACAAATCGGAACTGGAGGAACTACGTG	32
QY	101	ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla	12
Db	326	ACCCGCGTGGCGGAGAGACCGGGGCAACGCTGTCAAGAGAGCTGCACGCGCGCCAGGCC	38
QY	121	ArgLeuGlnYAlaAspMetClnAspValCysGlnYArgLeuValGlnTyrArgGlnGlnVal	14
Db	386	CGGCTGGCGCGGACATGAGAGAGCGTGGCGGCCCTGGTGCATGACCGCGGCGAGGTG	44
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg	16

Db	446	CAGGCGATGCTGGCGCAGAGCACCGAGGAGCTGGCGGTGGCGCTGCCACCTGGCG	505
OY	161	LysLeuCySLysAlaGLeuLeuAlaArgAlaAspAlaAspAlaGLeuLysAlaGLeuAlaValTyr	180
Db	506	AAGCGCTGAAGCGGCTCTCTCCGCATGCCGATGACCTGCAGAAACGGCTGGCACTGTAC	565
OY	181	GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeu-Gl	200
Db	566	CAGGCGGGGGCCGCGAGGGCGCGAGCGCGCTGCAGCGCCATCCGCGAGCGCTGGGG	625
OY	200	YProLeuValAlaGlnGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlnIlePr	220
Db	626	CCCCCTGGTGGACAGGCTCCGCGTCCGGGGCCGCCACCTGTGGCTCTCCCTGGCGGCGACCC	685
OY	220	oLeuGlnGluArgAlaGlnAlaIleTrpGlyGluArg-LeuAlaArgAlaArgMetGlnGlu-Met	239
Db	686	GCTACAGGAGCGGGCGCAAGGCTGGGGCGAGGGCGCTTGCGCCGCGGATGGAGGAAGAG	745
OY	240	GlySerArg-ThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlu-ValAlaArg	259
Db	746	GCGCAGCGGGGACCCGCGACCGCTGGACCGAGAGTGAAGAGCAGGTGGCGGAAGTGGCCGC	805
OY	259	IaLysLeu---GlnGlnAlaGlnGlnGln-IleArgLeuGln-AlaGlnAlaPheGlnIle	277
Db	806	CCCCAGCTGAGCAGGACCGAGGCCCAAGATACGCCCTGCAGGGCGGAGGCTCTCTAGGC	865
OY	277	aArgLeuLysSerTrpPheGluPro-----LeuValAlaLysPheGlnArgGlnIlePrAl	295
Db	866	CCGGCTCAGAGAGCTGGTTCGAACCCCTGGGTGGAAACAGACATGGCGGGCGCAATGGGGC	925
OY	295	aGlyLeuValGlnLysVal-GlnAlaAlaValAlaGlyThrSerAlaAlaProValProsera	315
Db	926	CGGCTGGTGGAAAAAGGTGCCAGGCTGCCGGGGGCGACCAACCGCCCTCTGGTGGCAGC-G	984
OY	315	spasnHis 317	
Db	985	ACAATCAT 992	

Location/Qualifiers

```

source
1. .842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4860585"
/clone_1b="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

```


Db 195 GGCACGGCTGGGAACCTGGCACTGGGTGCTTTTGGATTACCTGGCTGGGTGCAGACA 254
 QY LeusergluInuValgluInuLeuserSerInuValThrInuLeuArgAla 80
 Db 255 CTGTCTAGCAGAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGCTCACCAGAACACTAGGGCG 314
 QY LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
 Db 315 CTGATGACGACGACCATGAAGAGTGTGAAGGCTACAAATGGAACTGGAACTGGAAACAACACTG 374
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaGluAla 120
 Db 375 ACCCGGTGGGGAGAGACCGGGCAGCGCTGTCCAAGAGCTGCAGCGCGCCAGAGCC 434
 QY ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 Db 435 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCCCTGTGTCAGTACCGCGGAGAGTG 494
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 495 CAGGCATGCTCGGCCAGAGACCGAGAGAGCTGGGGTGGCGCTCCGCTCCACCTGGCGC 554
 QY 161 LysLeuGlyLysArgLeuLeuArgAlaAspAlaAspLeuGluLysArgLeuAlaValTyr 180
 Db 555 AAGCTGGGTAAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 614
 QY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGly 200
 Db 615 CAGGCGGGGGCGCGAGAGGGCGCGAGCGGCTCAGCGCATCCCGAGCGGCTGGGG 674
 QY 201 ProLeuValGluGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 675 CCCCTGGTGGAAACGGGCGCGCTGGCGGCCCACTGTGGGCTCCCTGGCGCGGACACCG 734
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArg-MetGlu---GluMe 239
 Db 735 CTACAGAGAGGGGCCAGGCGCTGGGGAGGGCTGGCGCGCGATGGAGGCACACTTG 794
 QY 239 TgLYSeArGThrArGAspArgLeuAsp-GluValLysGluGluAlaGluValArgA 259
 Db 795 GGGCAGCGGAGCCCGCAGACCGCTGGAACGAGTGAAGAGACAGTGGGAGAGTGGGCC 854
 QY 259 TalysLeuGluGluGlnAlaGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgL 279
 Db 855 GCCAAGCGTGGGGAGACCGAGCGCCCAATGCTTGCGAGGCCAA-GCCTTCCAGCGCGGCT 913
 QY 279 eulysSerTrpPheGlu 284
 Db 914 CCAAGAGCTGGTTCAAA 930

RESULT 4
 BG761746 938 bp mRNA linear EST 15-MAY-2001
 LOCUS 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4641411 5',
 DEFINITION mRNA sequence.
 ACCESSION BG761746
 VERSION BG761746.1 GI:14072399
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
 TITLE 1 (bases 1 to 938)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@nsf.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LUCM1674 row: c column: 12
 High quality sequence stop: 767.
 Location/Qualifiers

FEATURES
 source
 1. 938

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4641411"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCAGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using Zap-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

BASE COUNT 171 a 272 c 374 g 121 t
 ORIGIN

Alignment Scores:
 Pred. No.: 4.85e-117 Length: 938
 Score: 1212.00 Matches: 271
 Percent Similarity: 92.28% Conservative: 4
 Best Local Similarity: 90.94% Mismatches: 12
 Query Match: 76.27% Indels: 11
 DB: 12 Gaps: 2

US-09-827-854-18 (1-317) x BG761746 (1-938)

QY 1 MetLysValLeuThrPheAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 57 ATGAAGGTTCTGGGGCTGGCTGTGCTGTGCTACATCTCTGCGAGATCCAGGCCAAGAGTG 116
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 117 GAGCAACGGGTGGAGACAGACGCGAGACCGGAGCTGGCGCAGACGACGAGTGGAGAGC 176
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 177 GGCAGAGGCTGGGAACGTGACATGGCTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 236
 QY 61 LeuserGluInuValGlnGluLeuLeuserSerInuValThrInuLeuArgAla 80
 Db 237 CTGTCTAGCAGAGGTGCAGAGAGAGCTGTCAAGTCCAGGTCAACCCAGAACTGAGGGCG 296
 QY 81 LeuMetaspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
 Db 297 CTGATGACGACGACCATGAAGAGTGTGAAGGCTTACAAATCGGAACGTGAAGAACAACTG 356
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaGluAla 120
 Db 357 ACCCGGTGGCGGAGAGAGCGGGGCACGGCTGTGCCAAGAGCTGCAGCGCGCGAGAGCC 416
 QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 Db 417 CGGCTGGGCCCGGACAGAGAGAGCTGTGGCGCGCTGTGAGTGAAGTCCCGCGGAGAGTG 476
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 477 CAGGCATGCTCGGCCAGAGACCGAGAGAGCTTGGGTGGCTCCGCTTCCACTGCGCGC 536
 QY 161 LysLeuGlyLysArgLeuLeuArgAspAlaAspLeuGluLysArgLeuAlaValTyr 180
 Db 537 AAGCTGGGTAAAGGGGCTCCCGCGATGCCGATGACCTGCAGAAAGCGGCGGAGGTGATC 596
 QY 181 GlnAlaGlyAlaArgGluGluGluGluArgGlyLeuSerAlaIleArgGluArgGly 200
 Db 597 CAGGCGGGGCGCGAGAGGCGCGAGCGGCTCTCAGCGCCATCCGCGAGCGGCTGGGG 656

KEYWORDS	EST.
SOURCE	
ORGANISM	
	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 800)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

BASE COUNT	146 a	239 c	315 g	100 t
ORIGIN				

Alignment Scores:	
Pred. No.:	2,36e-115
Score:	1195.00
Percent Similarity:	98.80%
Best Local Similarity:	97.20%
Query Match:	75.20%
DB:	13
Length:	800
Matches:	243
Conservative:	4
Mismatches:	3
Indels:	1
Gaps:	0

RESULT 10
 B1600563
 LOCUS 811 bp mRNA 1 linear EST 07-SEP-2001
 DEFINITION 603244936F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287329 5',
 mRNA sequence.
 ACCESSION B1600563
 VERSION B1600563
 KEYWORDS GI:15493502
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 811)
 NIH-MGC http://mhc.nci.nih.gov/.
 NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L16M1726 row: g column: 10
 High quality sequence stop: 783.
 Location/Qualifiers
 1..811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5287329"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pBluescriptK1 (modified
 pBluescript K1); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 148 a 245 c 316 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.12e-114 Length: 811
 Score: 1181.00 Matches: 240
 Percent Similarity: 97.978 Conservative: 1
 Best Local Similarity: 97.568 Mismatches: 4
 Query Match: 74.328 Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-18 (1-317) x B1600563 (1-811)
 QY 1 MettlyValleutrpAlaalaLeuLeuValThrPheLeuAlaGlycysGlnAlaLysVal 20
 |||||||
 DB 74 ATGAAGTTCTGTGGCTGCTGTGGTGCATTCCTCGCAGAGATGCCAGGCCAAGTGC 133
 |||||||
 QY 21 GtugAlaValaGtutrhgrGtuprogLuprogLuleuArgInclntrhgrLutrpGInser 40
 |||||||
 DB 134 GACCAAGCGGTGAGACAGAGCCGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 193
 |||||||
 QY 41 GtGlnArGrppluLeuAlaLeuGtYArGpHeTrpAspTrpLeuArGrTrpValGIntrh 60
 |||||||
 DB 194 GGGCAGCGCTGGGAACTGGCAGCTGGGTTCGTTTGGGATTAAGTGGCTGGGTCAGACA 253
 |||||||
 QY 61 LeuSerGtugLInValGtIngtugLuleuLeuSerSerGlnValThrgIngtugLuleuArgAla 80
 |||||||
 DB 254 CTGTCTAGACAGGTGACGAGGAGGAGCTGTACCTCCAGGTACACCGAAGTGAAGGGCG 313

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
 |||||||
 DB 314 CTATAGACAGACACCATGAAGAGAGTTGAGAGCCTACCAATTCGAACTGGAGACAACTG 373
 |||||||
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 DB 374 ACCCGGTGGCGAGAGAGACGCGGCGAGCTGTCCAAAGAGAGCTGCAGGCGGCGAGGCC 433
 |||||||
 QY 121 ArgLeuGtYAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
 |||||||
 DB 434 CGGCTGGCGCCGACATGGAGACGTGTGGCGGCGCTGTGTGCATGCCCGGAGGTG 493
 |||||||
 QY 141 GtAlaMetLeuGtYgInserThrgIngtugLuleuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 494 CAGGCCATGCTGGCCAGACACCGAGAGACTCGGGTGGCGCTGCCCTCCACTGGCGC 553
 |||||||
 QY 161 LysLeuCySLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 554 AAGCTGGTAAAGGGGCTCTCCGCGATGCGCATGACCTGCAGAAAGCCCTGGCAGTGTAC 613
 |||||||
 QY 181 GtAlaGtYAlaArgGtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtug 200
 |||||||
 DB 614 CAGCGCGGGCGCCGAGAGGCGCGAGCGGCGCTCAGCCCATCCGCGAGCGCGCTGGGG 673
 |||||||
 QY 200 YProLeuValGtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtug 220
 |||||||
 DB 674 CCCCTTGTTGAGACAGAGCGCGCGCTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 733
 |||||||
 QY 220 OleuGtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtugLIngtug 240
 |||||||
 DB 734 GTACAGAGAGAGCGGCGCGAGCGCTGGGCGAGCGGCTGGCGCGCGCTGGAGCAATGGG 793
 |||||||
 QY 240 YSerArgTrpArgAsp 245
 |||||||
 DB 794 CAGCGGAGAACCGCGAA 809
 |||||||
 RESULT 11
 B6829472
 LOCUS 845 bp mRNA 1 linear EST 22-MAY-2001
 DEFINITION 602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:489112 5',
 mRNA sequence.
 ACCESSION B6829472
 VERSION B6829472.1 GI:14177059
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 845)
 NIH-MGC http://mhc.nci.nih.gov/.
 NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L16M1790 row: g column: 17
 High quality sequence stop: 829.
 Location/Qualifiers
 1..845
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:489112"
 /clone_lib="NIH_MGC_42"
 /tissue_type="epithelioid carcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOMB7; Site:1: XhoI;

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 253 CTGCTGAGAGAGGTGACGAGGAGGAGCTGCTCCAGCTCCAGGTCACGAGAGACGAGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||
 Db 313 CTGATGAGAGACCATGTAAGGAGTGAAGGCTTACAAATCCGAACTGGAGGAGAACTG 372
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 373 ACCCGGTGACGAGGAGACGCGGCGCAGCTCTCCAGAGAGCTGCAGCGGCGGAGGCC 432
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 |||||
 Db 433 CGGCTGGGCGGACATGAGAGACGCTGCGCGGCTGTGTGACAGTACCGCGGAGGTC 492
 QY 141 GlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 493 CAGGCAATGCTCGCGCAGAGCACCGAGAGCTCGGGTGGCTCGCTCCCTCCACTGGCG 552
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 553 AACCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTCGAGAGCGCTGGCAGGTAC 612
 QY 181 GlnAlaGlyAlaArgGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 200
 |||||
 Db 613 CAGCGCGGCGCGCGAGGCGCGCGAGCGGCTCTCAGCGCCCATCCGCGGCGCGCTGGGG 672
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||
 Db 673 CCCCTGTGGAACAGGCGCGCTGCGCGCGCATGTGGGCTCCCTGGCGACGACGCG 732
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGlnMet 239
 |||||
 Db 733 CTACAGAGAGCGGCGCGAGGCTGGGCGCGAGCGGTG-CGCGCGCGGATGAGAGATG 788
 RESULT 13
 B1551475 919 bp mRNA linear EST 05-SEP-2001
 LOCUS B1551475
 DEFINITION 603194314P1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
 mRNA sequence.
 ACCESSION B1551475
 VERSION B1551475.1 GI:15438787
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 919)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NMGRI), Shiraki
 Toshitaki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11691 row: 1 column: 04
 High quality sequence stop: 812.
 Location/Qualifiers
 1..919
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5274003"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTCTTTTCTTTTCTTTT-3',
 size-selected for average insert size 2.5 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIMH/NMGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 172 a 270 c 363 g 114 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,37e-109 Length: 919
 Score: 1141.00 Matches: 246
 Percent Similarity: 93.26% Conservative: 17
 Best Local Similarity: 92.13% Mismatches: 17
 Query Match: 71.81% Indels: 4
 DB: 13 Gaps: 0
 US-09-827-854-18 (1-317) x B1551475 (1-919)
 QY 1 MetLysValLeuThrAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 73 ATGAAGTTCTGTGGGCTGGCTGGTGTGTCACATCTCTGGCAGAGATGCCAGGCAAGTGC 132
 QY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTyrPglIns 40
 |||||
 Db 133 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGCCAGAGACCGAGTGGCAGAGC 192
 QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrPValGlnThr 60
 |||||
 Db 193 GGCACGCTGGGAACTGCGCAGTGGCTCTTTGGGATTAAGTGGCTGGTGGTGGAGACA 252
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 253 CTGCTGAGAGAGGTGACGAGAGAGTGTGCTCAGCTCCAGGTACCCAGGAATGAGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||
 Db 313 CTGATGAGAGACCATGTAAGGAGTGAAGGCTTACAAATCCGAACTGGAGGAACAATG 372
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 373 ACCCGGTGACGAGGAGAGACGCGGCGCAGGCTCTCCAGAGAGCTGCAGCGGCGGAGGCC 432
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 |||||
 Db 433 CGGCTGGGCGGACATGAGAGACGCTGCGCGGCTGTGTGACAGTACCGCGGAGGTC 492
 QY 141 GlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 493 CAGGCGGCGGCGCGAGTGGCGCGGAGCTGTGCGCGCATCCGCGAGCGCTGGGG 672
 QY 161 LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 553 AACCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTCGAGAGCGCTGGCAGGTAC 612
 QY 181 GlnAlaGlyAlaArgGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 200
 |||||
 Db 613 CAGGCGGCGGCGCGAGTGGCGCGGAGCTGTGCGCGCATCCGCGAGCGCTGGGG 672
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 |||||
 Db 673 CCCCTGTGGAACAGGCGCGCTGCGCGCGCATGTGGGCTCCCTGGGCGGACGCCCT 732
 QY 221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluGlnMet 240
 |||||
 Db 733 A-CAGAGACGCGGCGCAGGCTGGCGGAGCGGTGGCGCGGATGAGAGGAATGGG 790
 QY 240 ySerArgThrArgAspArgLeuAspGluValLysGluGlnAlaGluValArgAlaLys 260
 |||||
 Db 791 CAGCCGAGCCGCGGAGCG--CTGGACGAAGTGAAGACACGCTGGCGGAGAGG-TGGCCCAA 847

Plate: LLCM2021 row: k column: 21
High quality sequence stop: 757.

FEATURES	Location/Qualifiers
source	1. .757

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/clone_id="NIH_MGC_112"
/tissue_type="melanocytic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

Alignment Scores:	
Pred. No.:	1.16e-108
Score:	131.00
Percent Similarity:	98.73%
Best Local Similarity:	98.73%
Query Match:	71.18%
DB:	13
Length:	757
Matches:	234
Conservative:	0
Mismatches:	2
Indels:	0
Gaps:	0

US-09-827-854-18 (1-317) x BM042228 (1-757)

QY 1 MetllysvallLeutrpalaalaLeuValThrphLeuValaglyCysGlnAlaLysVal 20

DB 30 AAGAGGTCCTGTCGGCTGCGTGTCTGGTCACATTCCTGGCAGGATGCCAGGCCCAAGGTG 109

Case	Age	Sex	Occupation	Duration of illness	Site of lesion	Pathological changes	Microscopic findings	Diagnosis
1	45	M	Farmer	10 years	Right lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer
2	55	F	Housewife	5 years	Left lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer
3	65	M	Farmer	15 years	Right lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer
4	75	F	Housewife	20 years	Left lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer
5	85	M	Farmer	25 years	Right lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer
6	95	F	Housewife	30 years	Left lower leg	Chronic ulcer	Ulcer with fibrinous exudate, necrotic debris, and inflammatory cell infiltration.	Chronic ulcer

Qy	61	LeSerGIuGIuValGIuGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgIa	80
Db	230	CTGTGTGACGACAGTCCAGAGAGAGCTGCTCAGCTCCACAGTACCCAGAGACTGAGGGCG	289
Qy	81	LeuMetAspGIuThrMetIysGIuLeuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu	100
Db	290	CTGATGGACGAGACCATGAGAGATTGAAGGCTTCACAAATCGCAACTGAGAGACCACTG	349
Qy	101	ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaAlaGIuAla	120
Db	350	ACCCGGGTGGCGAGAACGCGGGCGACGGCTGTCCAAAGGACTTCACAGCGGCGCAGGGCC	409
Qy	121	ArgLeuGIuAlaAspMetGIuAspLysAlaLysGIuArgLeuValGIuTyrArgGIuLysAla	140
Db	410	CGGGTGGCGCGGACATGAGAGACTGTGGCGGCCCTGTGGTCCACTACCGCGGCAGGGT	469
Qy	141	GlnAlaMetLeuGIuGlnSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	470	CAGGCGATGCTGTGGCCAGAGCACCGAGAGAGCTGGGGGTGGCGCTGTGCTCCACCTGGGC	529
Qy	161	LysLeuCysLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr	180

Db 650 CCCCTGGTGAGACAGGGCCGGCGTTGGCGGGCCGCCACATGTGGGCTCCCTGGCGCGGCCAACC 709

Db	710	GCTACAGGACGGCGGC-CAGGCCTGGGGGAGAGGCTCGCGGGCGGATG	757
RESULT 15			
LOCUS	BG706129	907 bp	EST 07-MAY-2001
DEFINITION	BG706129	602668093F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5', mRNA sequence.	
ACCESSION	BG706129		
VERSION	BG706129.1	GI:13981169	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 907)		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .		

JOURNAL COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapps-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

Clone distribution: MGC clone distr

<http://image.1lnl.gov/photo/RAW10660.jpg> .

High quality sequence stop: 83.2.
Location/Qualifiers

```
source      1.1307
/organism="Homo sapiens"
```

```
/tissue_type="hypothalamus"
/lab_host="DH10B"
```


GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2307.23 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-19
Perfect score: 1585
Sequence: 1 MKVLMAALLVFLACQAAKV.....VEKVOAAVGSAAVPSPDNH 317

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=rlp
-Q=/cgn2.1/USPRO.spool/US09827854/rluat.11032003.101610.27486/app.query.fasta.1.3576
-DB=GenBdl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEADSIZE=5001 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854 @CGN.1.1.13008 @runat.11032003.101610.27486 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEJUNKER -NEGSCORES=0 -WAIT -LONGIO -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBdl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_in:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1581	99.7	1110	6	E00359	E00359 cDNA coding
2	1581	99.7	1110	6	E00823	E00823 DNA sequence
3	1581	99.7	1147	6	AX302545	AX302545 Sequence
4	1581	99.7	1156	6	BD004278	BD004278 Apo E hum
5	1581	99.7	1156	9	HUMAP0E3	K00396 Homo sapien
6	1581	99.7	1186	9	BC003557	BC003557 Homo sapi
7	1574	99.3	1110	6	E08423	E08423 DNA coding
8	1573	99.2	1156	6	BD004277	BD004277 Apo E hum
9	1572	99.2	1157	6	AX333278	AX333278 Sequence
10	1572	99.2	1157	6	AX409597	AX409597 Sequence
11	1572	99.2	1157	6	I15975	I15975 Sequence 1
12	1572	99.2	1157	9	HUMAP0E	M12529 Human apoli
13	1569	99.0	1156	6	BD004279	BD004279 Apo E hum
14	1447	91.3	1108	9	MFAPOE	X13887 Monkey mRNA
15	1392.5	87.9	5491	9	AF261279	AF261279 Homo sapi
16	1392.5	87.9	5491	6	AX358722	AX358722 Sequence
17	1392.5	87.9	41907	9	AF050154	AF050154 Homo sapi
18	1392.5	87.9	107567	9	AC011481	AC011481 Homo sapi
19	1375	86.8	5515	9	HUMAP0E4	M10065 Human apoli
20	1366	84.9	5413	9	AF261280	AF261280 Pan trogl
21	1263	79.7	4762	9	BABAP0E	M29322 Baboon apoli
22	1262	79.6	208339	2	AC021988	AC021988 Homo sapi
23	1254	79.1	1138	4	AF303830	AF303830 Tupaiat gl
24	1166	73.6	1060	4	RABAP0LP	M36603 Rabbit apoli
25	1159.5	73.2	965	6	AX384545	AX384545 Sequence
26	1159.5	73.2	1108	4	BTAPOLE	X61171 B.taurus mr
27	1159.5	73.2	5617	6	AX384541	AX384541 Sequence
28	1159.5	73.2	6026	6	AX384539	AX384539 Sequence
29	1147	72.4	718	9	AF20049753	AF200497 Pan trogl
30	1145	72.2	1045	10	MUSAP0E	M12414 Mouse apoli
31	1145	72.2	1104	10	BC028816	BC028816 Mus muscu
32	1143	72.1	718	9	AF20050053	AF200500 Gorilla g
33	1143	72.1	718	9	AF20050353	AF200503 Pongo pyg
34	1141.5	72.0	1154	9	AF20050653	AF200506 Hylobates
35	1139	71.9	718	9	AF20050653	AF200506 B.taurus mr
36	1139	71.9	1122	4	SSAP0E	X72835 S.scrofa mr
37	1134	71.5	1126	6	ARI64342	ARI64342 Sequence
38	1134	71.5	1126	6	ARI64342	ARI64342 Sequence
39	1090	68.8	951	10	AF205885	AF205885 Sequence
40	1068	67.4	959	10	MUSAP0E	M73490 Mus musculu
41	1052.5	66.4	1069	10	RATAP0E	J00705 rat apoli
42	1025	64.7	228698	2	AC127479	AC127479 Mus muscu
43	1025	64.7	237653	2	AC073760	AC073760 Mus muscu
44	999	63.0	4856	10	MUSAP0E	D00466 Mus musculu
45	998	63.0	4267	6	ARI64387	ARI64387 Sequence

RESULT 1

ALIGNMENTS

E00359
LOCUS E00359 1110 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA coding human apolipoprotein E3.
ACCESSION E00359
VERSION E00359.1 GI:2168646
KEYWORDS JP 1985118189-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1110)
AUTHORS Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.
TITLE DNA FRAGMENT
JOURNAL Patent: JP 1985118189-A 1 25-JUN-1985;
MITSUBISHI CHEM IND LTD
COMMENT OS human JP 1985118189-A/1
PN 25-JUN-1985
PD 29-NOV-1983 JP 1983224980
PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
KIMURA MASAKO,
PI IKEDA YASUOKO
PC C12N15/00.C07H21/04//C12P21/00;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue: liver;
FH key Location/Qualifiers
FT CDS 15..968
FT /product='apolipoprotein E3 precursor' FT
FT sig_peptide 15..68
FT /product='apolipoprotein E3 signal peptide' FT
FT mat_peptide 69..965
FT /product='apolipoprotein E3'
FEATURES
Source Location/Qualifiers
1..1110
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 5.65e-80 Length: 1110
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-19 (1-317) x E00359 (1-1110)
QY 1 MetLysValLeuThrPalaIalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 15 ATGAAGGTTCCTGTCGGTGGCTGCTGCACATTCCTGGCAGATGCCAGGCCAAGTG 74
QY 21 GluAlaAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 75 GAGCAACCGGTGGAGACAGACCCGAGCTGGCCACGACGACGAGCTGGCAAGC 134
QY 41 GlyLnaIArgTrpGluLeuAlaIalaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 135 GGCAGCGCTGGAGAACCTGGCAGCTGGCTGTTGGGATTACCTGGCTGGGTGCACACA 194
QY 61 LeuSerLgluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 195 CTGTCTAGAGAGGTGCGAGGAGGAGCTCTCAGCTCCAGGTCACCCAGGAAGTGAAGGCG 254
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
Db 255 CTGATGACGACGACCATGAAGGAGCTTAAGGCCCTACAAATCGAAGCTGAGGAACAACCTG 314

QY 101 ThrProValAlaGluGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 315 ACCCGGTGGCGGAGAGACCGCGGGCAGCGCTGTCCAAAGAGCTGCAGCGCGCGCAGAGCC 374
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 375 CGGCTGGGCGCGACATGAGAGACGTGTGCGCCGCTGTGCATGTACCGCGGAGAGGTG 434
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGCGCATGCTGGGCCAGAGACGACGAGAGAGTGGGGTGGCCCTCGGCTCCACACTGCC 494
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 495 AAGCTGGTAAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCCCTGGCAGTGTAC 554
QY 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgLysLeuSerAlaIleArgGluArgLeuGly 200
Db 555 CAGCGCGGGCGCGCGAGGGCGCGAGCGCGGCTCCAGCGCATCCCGCGAGCGCTGGGG 614
QY 201 ProLeuValGluGlnArgValArgValArgAlaAlaThrValGlySerLeuAlaGlyPro 220
Db 615 CCCCTGGTGGACAGGCGCGCGTGCAGCGCGCCACTGTGGCTCCCTGGCCGCGCAGCG 674
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 675 CTACAGAGAGCGGCGCCAGCGCTGGGGCGAGCGGCTGCGCGCGGATGAGAGATGGCG 724
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 735 AGCGGAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 794
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 795 CTGAGAGAGCAGCGCCACGACATGCGCTGCGAGCGCAGGCGCTTCACGCGCCGCTAAG 854
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 855 AGCTGTTTCAGAGCCCTCGTGTGAAGACATGCAGCGCCAGTGGCGCGGTGTGAGAGAG 914
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 915 GTGCAGGCTCGCGTGGCACCGCGCCGCTGTGTGCCAGCGACAATCATC 965
RESULT 2
E00823 1110 bp DNA linear PAT 29-SEP-1997
LOCUS E00823
DEFINITION DNA sequence coding for human apolipoprotein E and its signal
peptide.
ACCESSION E00823
VERSION E00823.1 GI:2169084
KEYWORDS JP 1986096997-A/1.
SOURCE unidentifed.
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 1110)
AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
MITSUBISHI CHEM IND LTD
COMMENT OS Human (Homo sapiens)
PN JP 1986096997-A/1
PD 15-MAY-1986
PI 16-OCT-1984 JP 1984216987
PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO
PC C12P21/00.A61K37/4.A61K37/04.C12N15/00.(C12P21/00,C12R1:19),
PC (C12N15/00,
PC C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue: liver;

Key Location/Qualifiers
FH 3'UTR 1..14
FT sig-peptide 15..68
FT /product='human apolipoprotein E signal' FT
FT CDS peptide: 69..968
FT /product='human apolipoprotein E' FT 3'UTR
FEATURES
source location/Qualifiers
1..1110
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 198 a 353 c 416 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 5.65e-80 Length: 1110
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: Gaps: 0
US-09-827-854-19 (1-317) x E00823 (1-1110)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAGTTCTGTGGCTGCGCTGCTGTCACATTCTTGCGAGATGCCAGGCCAAGTG 74
QY 21 GlnGlnAlaValAlaGlnTrpGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
DB 75 GAGCAACGCGTGGAGACAGAGCCGAGCCGAGCTGGCCACAGACAGACGAGTGGCAAGC 134
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 135 GGCAGAGCGTGGAACTGGCACTGGCTTTGGGATTACCTGGCTGGTGGCAAGCA 194
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 195 CTGCTCAGACGAGTGGAGAGAGAGCTCTCAGCTCCAGGTCACCCAGAACTGAGGGCG 254
QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGlnLeuGlnGlnLeu 100
DB 255 CTGATGACGAGACCATGAAGAGATTGAAGCCCTACAAATCGGAACTGGAGAACACTG 314
QY 101 ThrProValAlaGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 315 ACCCGGTGGCGGAGAGAGCGGGCGCTGTCCAAGAGAGTGCAGCGCGGCGCAGGCC 374
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
DB 375 CGGCTGGGCGGAGCATGAGAGAGAGTGTGGCGCGCTGTGGCTGACATGCCGCGAGGTG 434
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGGCGCAGACACGAGAGAGTGGGGTGGCGCTCCACTCCAGCTGCC 494
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
DB 495 AAGGTGGTAAAGGGGCTCTCCGCGATGCCGATGCAAGAGCCCTGGCAGTTTAC 554
QY 181 GlnAlaGlnAlaArgGlnGlnValGlnArgGlyLeuSerLysAlaGlnArgLeuGly 200
DB 555 CAGGCCGCGGCGCGGAGGCGCGAGCGCGGCTCAGCCCAATCCGCGAGCGCTGGGG 614
QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValLysSerLeuAlaGlnPro 220
DB 615 CCCGTGTGAACAGGCGCGCTGCCGCCCACTGTGGGCTCCCTGGCGCGGCGCAGCCG 674
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 675 CTACAGAGAGCGGCGCGGCTGGGGGAGCGGCTGCCGCGGATGAGAGATGGCG 734

QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 735 AGCCGAGCCCGGACCCCTGGAGAGGTGAAGAGACAGAGTGGCGAGGTGCCGCCAAG 794
QY 261 LeuGlnGlnAlaGlnGlnAlaArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 795 CTGGAGAGAGAGGCGCAGAGATACGCTCGAGGCGCAGAGCCTTCCAGGCCCGCTCAG 854
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 855 AGCTGTTTCAGAGCCCTGGTGAAGACATGCAGCGCCAGTGGCGGCTGTGGAGAG 914
QY 301 ValGlnAlaAlaValAlaGlnTrpSerAlaAlaProValProSerAspAsnHis 317
DB 915 GTCCAGGCTGCTGGGCGACAGCGCGCCCTGTGTGCCAGCCAAATTCAC 965
RESULT 3
AX302545 1147 bp DNA linear PAT 30-NOV-2001
LOCUS AX302545
DEFINITION Sequence 63 from Patent WO0175177.
ACCESSION AX302545
VERSION AX302545.1 GI:17383082
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Morin,P.J., Sherman-Baust,C.A., Pizer,E.S. and Hough,C.D.
Tumor markers in ovarian cancer
Patent: WO 0175177-A 63 11-OCT-2001;
JOURNAL THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES
source location/Qualifiers
1..1147
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 210 a 365 c 425 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 5.84e-80 Length: 1147
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: Gaps: 0
US-09-827-854-19 (1-317) x AX302545 (1-1147)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGTTCTGTGGCTGCGCTGCTGTCACATTCTTGCGAGATGCCAGGCCAAGGTG 105
QY 21 GlnGlnAlaValAlaGlnTrpGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
DB 106 GAGCAACGCGTGGAGACAGAGCCGAGAGCTGGCCACAGACAGACGAGTGGCAGAGC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 166 GGCAGAGCGTGGAACTGGCACTGGTGGCTTTGGGATTACCTGGCTGGTGGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 226 CTGCTGAGAGAGTGCAGGAGAGAGTGTCTTCCAGGTCACCCAGGAACCTGAGGGCG 285
QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGlnLeuGlnGlnLeu 100
DB 286 CTGATGACGAGACCATGAAGAGATTGAAGCCCTACAAATCGGAACTGGAGAACACTG 345
QY 101 ThrProValAlaGlnGlnTrpArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 346 ACCCGGTGGCGGAGAGAGAGCGGGCGCTGTCCAAGAGAGCTGCAGCGCGCGCAGGCC 405

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 406 CGCGTGGGCGGACATGAGAGAGCTGTGGCGCGCTGTGGTGCAGTACGGCGGAGGTG 465
 |||||||
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 466 CAGCCCAATGCTCGCCCAAGACAGCAGAGAGCTGGGGGTGGCTCGCTCCACCTGGCG 525
 |||||||
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 526 AAGCTGCGTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAACCGCTTGCACTGTAC 585
 |||||||
 QY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 Db 586 CAGCGCGGGCGCCGCGAGCGCGCGAGCGCTCCACGCGCATCCGCGAGCGCTTGCGG 645
 |||||||
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 Db 646 CCCCCTGGTGAACAGGGCGCGGTGGGGCGCCCACTGTGGCTCCCTGGCGCGCGCAAG 705
 |||||||
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 |||||||
 Db 706 CTACAGAGAGCGGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC 765
 |||||||
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||||
 Db 766 AGCCGAGACCCCGACCGCTGTGAGAGGTGAGAGCAGTGGCGGAGAGTGGCGCCCAAG 825
 |||||||
 QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 |||||||
 Db 826 CTGGAGGGAGCGAGCCGACAGATACGCTTGAGCGCGAGCGCTTCCAGCCCGCTCAAG 885
 |||||||
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||||
 Db 886 AGCTGCTTCGAGCCCTCTGTGGAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 945
 |||||||
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||||
 Db 946 GTGCAGGCTGCCGTGGGACACAGCGCCCGCTGTGCCACAGCAATCAC 996
 |||||||
 RESULT 4
 BD004278 1156 bp DNA linear PAT 31-JAN-2002
 LOCUS BD004278 Apo E humanized mammal.
 DEFINITION BD004278
 ACCESSION BD004278.1 GI:18632239
 VERSION BD004278.1 GI:18632239
 KEYWORDS JP 2001017028-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1156)
 Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
 Apo E humanized mammal
 Patent: JP 2001017028-A 2 23-JAN-2001;
 MITSUBISHI CHEMICAL CORP
 OS Homo sapiens (human)
 PN JP 2001017028-A/2
 PD 23-JAN-2001
 PR 28-APR-2000 JP 2000128919
 FH SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
 A01K67/027, A61K45/00, A61P25/28, A61P43/00, C12N5/10, PC
 C12N5/09//C07K14.775,
 PC (C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC
 FEATURES
 source FH Key Location/Qualifiers
 FT CDS (61)..(1011).
 1..1156 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 208 a 368 c 432 g 148 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,89e-80 Length: 1156
 Score: 1581.00 Matches: 316
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.68% Mismatches: 0
 Query Match: 99.75% Indels: 0
 DB: Gaps: 0
 US-09-827-854-19 (1-317) x BD004278 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 61 ATGAAGGTTCTGTGGGCTGCGCTGTGCTGATTCCTGCGCAGAGATGCAGGCCAAGGTG 120
 |||||||
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||||
 Db 121 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGCCCGCCAGACAGACGAGTGCAGAGC 180
 |||||||
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 Db 181 GGCCAGCGCTGGAACTGGCACTGGGTGCTTTTGGATTAACCTGCGCTGGGTGCAGACA 240
 |||||||
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 Db 241 CTGTCGTAGCAGGTGACAGAGAGAGCTGCTCACTGCCAGCTCCACAGCACTAGAGGCG 300
 |||||||
 QY 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
 |||||||
 Db 301 CTGATGAGCAGAGACCATTAAGAGAGTTGAAGCCCTCAATGGAATCGAGAACTGAGAACTG 360
 |||||||
 QY 101 ThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||||
 Db 361 ACCCGGTGGGGAGAGAGAACCGCGGCAACGCTGTCCAAAGAGAGCTGCAGAGCGCGCAGGCC 420
 |||||||
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 421 CGCGTGGCGCGGACATGAGAGAGGTGTGGCGCGCGCTGTGCGAGTACCCGCGGAGAGTG 480
 |||||||
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 481 CAGCGCATGCTGGCCCAAGAGCAGCAGAGAGCTGGGGCGCTTCGCTCCACCTGGCG 540
 |||||||
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 541 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCACTGTAC 600
 |||||||
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 Db 601 CAGCGCGGGCGCCGAGGGCGCGAGCGCGCTCCACGCGCATCCGAGCGGCTGGGG 660
 |||||||
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 Db 661 CCCCCTGTGGAACAGGGCGCGCTGGGGCGCCCACTGTGGCTCTCCCTGGCGCGCAGCG 720
 |||||||
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 |||||||
 Db 721 CTACAGAGAGCGGGCCAGGCTGTGGGGGAGCGGTGGCGCGCGGATGAGAGATGGGC 780
 |||||||
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 |||||||
 Db 841 CTGGAGAGAGCGGCCAGCAGATAGCTTGAGGCGCAGGCGCTTCCAGGGCGCGCTCAAG 900
 |||||||
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 |||||||
 Db 901 AGCTGTTTCAGACCCCTGTGTGGAGACATGCAGCGCCCAAGTGGGCGGCTGTGGAGAGAG 960
 |||||||

RESULT 5	HUMAPOE3	1156 bp	mRNA	linear	PRI 24-NOV-2000
LOCUS	HUMAPOE3				
DEFINITION	Homo sapiens preapolipoprotein E (APOE) mRNA, complete cds.				
ACCESSION	K00396				
VERSION	K00396.1				
KEYWORDS	apolipoprotein; apolipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 355 to 1156) Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W., Lofquist-Kahl,F., Karathanasis,S.K. and Zannis,V.I.				
TITLE	Identification and DNA sequence of a human apolipoprotein E cDNA clone				
JOURNAL	J. Biol. Chem.	257 (24)		14639-14641	(1982)
MEDLINE	83082756				
PUBMED	6897404				
REFERENCE	2 (bases 250 to 777) Wallis,S.C., Rogné,S., Gill,L., Markham,A., Edge,M., Woods,D., Williamson,R., and Humphries,S.				
AUTHORS	The isolation of cDNA clones for human apolipoprotein E and the detection of APOE RNA in hepatic and extra-hepatic tissues				
TITLE	EMBO J. 2 (12), 2369-2373				(1983)
JOURNAL	84131952				
MEDLINE	6199196				
PUBMED					
REFERENCE	3 (bases 1 to 1156) Zannis,V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.				
AUTHORS	Synthesis, intracellular processing, and signal peptide of human apolipoprotein E				
TITLE	J. Biol. Chem.	259 (9)		5495-5499	(1984)
JOURNAL	84185684				
MEDLINE	6325438				
PUBMED					
REFERENCE	4 (bases 88 to 1156) McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.				
AUTHORS	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant				
TITLE	J. Biol. Chem.	259 (10)		6498-6504	(1984)
JOURNAL	84212473				
MEDLINE	6327682				
PUBMED					
REFERENCE	5 (bases 577 to 624) Gill,L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N.				
AUTHORS	Isolation and characterisation of a variant allele of the gene for human apolipoprotein E				
TITLE	Biochem. Biophys. Res. Commun.	130 (3)		1261-1266	(1985)
JOURNAL	85279526				
MEDLINE	2992507				
PUBMED					
REFERENCE	6 (sites) Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P.				
AUTHORS	type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant				
TITLE	J. Clin. Invest.	83 (4)		1095-1101	(1989)
JOURNAL	89198059				
MEDLINE	2539388				
PUBMED					
COMMENT	[1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant. [5] epsilon-2 allele. [6] sites; mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4,				

epsilon-3 and epsilon-2. This sequence appears to be of the
 epsilon-3 allele. [1] argues that the apo E polymorphism involves
 mutations in the structural coding region; for example the
 epsilon-2 phenotype which is characterized by hyperlipoproteinemia
 is thought to result from a c to t change (arg to cys) at base 586
 below [3],[5]. The sequence shown is 57% homologous with human apo
 A-I and 81% homologous with rat apo E. For the epsilon-4 sequence,
 see the separate entry
 [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised
 relative to [1] to record the
 revisions in the FEATURES table. The sequence below is that of [J.
 Biol. Chem. 258, 11422-11422 (1983)]
 and [3]:
 Apo E is located on chromosome 19 --Jackson, Bruns and Breslow,
 PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to
 the Apo C-II gene (see separate entry).
 The two mutations causing type III hyperlipoproteinemia
 (apolipoprotein E phenotype E3/3) produces substitutions of Arg for
 Cys at amino acid 112 and Cys for Arg at amino acid position 142.
 Draft entry and printed copy of sequence for [1] kindly provided by
 L.L. Gill, University of Aberdeen.
 Complete source information:
 Human liver [1], [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3],
 [4] and blood [5], cDNA to mRNA.
 Location/Qualifiers
 1..1156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /tissue_type="liver and blood"
 1..1156
 /gene="APOE"
 61..1014
 /gene="APOE"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59546.1"
 /db_xref="GI:178851"
 /db_xref="GDB:G00-119-691"
 /translation="MKVIMAAALVTPLFAGCOAKVEAOAVETPEPELROTFMOSGGRMK
 EIALGRFMDYLRMVQTLSEFOVEELSSOYTOELRALMDETMKELKAYKSELROLPP
 VAEETRRLSKLELOAAORLGAQDMEDYCGRLVYVRGEVQAMLGOSTELRLASHLR
 KLRRLRLADADLDQRLAVYQAGARBEAENGSLAIRRLGPVLEQGVRAATVGSILAQ
 QPDLERQAQNGEELRLARMEDEGSRTRRLDEVKQVAVETRAKLEEDQAQIIRLQAEAFQ
 ARLSNFEPLVEVDQRQWAGIYERKVAVGTSAAPVPSDNH"
 61..114
 /gene="APOE"
 /note="apolipoprotein E"
 115..1011
 /gene="APOE"
 /product="apolipoprotein E mature peptide"
 267
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3],
 [4]"
 /catalation=[2]
 /replace="c"
 279
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
 /catalation=[4]
 /replace="a"
 288
 /gene="APOE"
 /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)]. [3]"
 /catalation=[4]
 /replace="a"
 369
 /gene="APOE"
 /note="in variant"
 /catalation=[4]
 /replace="a"
 409

Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN.ac: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: n Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557324.

FEATURES

source

1. 1186

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:1571 IMAGE:3355712"
/issue_type="Eye, retinoblastoma"
/clone_lib="NIH-MGC-16"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
46. .999
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAH03557.1"
/db_xref="GI:18097699"

CDS

/translation="MKLVAAALVTFLLACQAKVEQAVTEPEPELRQOTEMSGQRM
ELALGFMDYLRWVOTLSBOVEELSSVQTELRMLMETKAYSELEQITP
VAETFRARLSKELOAARLGAMDEVCRLVQKREVOAMLGOSTEELRSLASHLR
KLRLRLRDADLOKRLAYOAGAREGASRLAISREGLVDEGRVAAVYSLAG
OLOEBAOAMGERELRAMEMGRTRDRDEKVEQAVFRATLEDAQOIRLOAEAFQ
ALKSKPEELVEDMDKORWAGLVKYOAGVTSAAAPSPDNH"

BASE COUNT 248 a 366 c 425 g 147 t
ORIGIN

Alignment Scores:

Pred. No.: 6,05e-80 Length: 1186
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
Gaps: 0

US-09-827-854-19 (1-317) x BC003557 (1-1186)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAGGTTCTGTGGCTGTGGCTGTGCATTCCTCTGCGCAGAGTCCAGGCCAAGGTG 105
QY 21 GlnGlnAlaValAlaLthrGluProGluLthrGluLeuArgGlnGlnLthrGluProGlnSer 40
DB 106 GACCAAGCGGTGGAGACAGGCGGAGCCGACCTGCGCAGCAGACCGAGTGGCAGAGC 165
QY 41 GlyGlnArgTrpLuleuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 166 GGCAGAGCGGTGGAGACCTGGCTTTTGGATTACTCGCTGGTGGTGCAGACA 225
QY 61 LeuSerGlnGlnValGlnGluLuleuLeuSerSerGlnValThrGlnGlnLuleuArgAla 80
DB 226 CTGTCTGAGCAGCTGACGAGGAGCTCAGCTCCAGCTCCAGCAGCACTGAGGCG 285
QY 81 LeuMetAspGlnThrMetLysGlnLuleuLysAlaTyrLysSerGlnLuleuGlnGlnLuleu 100
DB 286 CTATGTGACAGACCATGAAGAGATTGACGCTTACAAATCGGAACCTGGAGAACACTG 345
QY 101 ThrProValAlaGlnGluLthrArgAlaArgLeuSerLysGlnLuleuGlnAlaAlaGlnAla 120
DB 346 ACCCGGTGGCGAGAGACGCGGCTGTCCAGAGAGCTGCAGCGGGCGCAGGCC 405
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140

DB 406 CGCGTGGCGCGGACATGAGAGACGTGTGGCGCCCGCTGTGTCAGTACCGCGGCGAGGTG 465
QY 141 GlnAlaMetLeuGlnGlnInsThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 466 CAGGCACTGCTCGGCGAGACACCGAGAGCTCGGGTGGCTTCCCTCCACTTCCGCG 525
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 526 AACCTGGTAAAGGGGCTCCCGCGATGCGGATGACCTGCAAGAGGCGCTGGCAGGTAC 585
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 586 CAGCGCGGCGCGCGCAGAGCGCGCGAGCGCGCTCAGCCCATCCCGCAGCGCTGGGG 645
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 646 CCCCTGTGGAACAGGGCGCGGTGGCGCGCGCTGTGGCTTCCCTGGCGCGCGCAGCG 705
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
DB 706 CTACAGAGAGCGGGCCCGAGCGCTGGCGGAGCGCTGGCGCGGATGAGAGATGGC 765
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys 260
DB 766 AGCGGAGACCGCGACCGCTGAGCAGAGTGAAGAGACAGTGGCGGAGTGGCGCCAG 825
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaGlnAlaArgLeuLys 280
DB 826 CTGAGAGAGAGCGCCAGCAGATACGCTGACGCGCGAGCGCGCTCCAGCGCGCTCAG 885
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB 886 AGCTGTGTGACGCCCTGCTGTGGAAGATCATCAGCCAGTGGCGCGCGCTGTGGAGAG 945
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTCCAGAGCTGCTGCTGGGACACAGCGCGCGCTGTGCTCCAGCAGCAATACAG 996
RESULT 7
E08423 1110 bp RNA linear PAT 29-SEP-1997
LOCUS
DEFINITION DNA coding human apolipoprotein E3.
E08423
ACCESSION E08423.1 GI:2176540
VERSION JP 1994315392-A.1
KEYWORDS Homo sapiens.
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Morimoto,H. and Teranishi,Y.
TITLE METHOD FOR PRODUCING APOLIPOPROTEIN
JOURNAL Patent: JP 1994315392-A.1 15-NOV-1994;
MITSUBISHI KASEI CORP
COMMENT
OS Homo sapiens (human)
PN JP 1994315392-A/1
PF 11-JUN-1985 JP 1994015433
PE 15-NOV-1994
PT MORIMOTO HIRONORI, TERANISHI YUTAKA
PC C12P21/02,C07K13/00,C12N5/10,C12N15/12,C12N15/18,C12P21/02,
PC C12R1.91),
PC (C12N5/10,C12R1.91);
CC strandedness: Double;
CC topology: Linear;
FH key
FH Location/Qualifiers
FT source 1. 1110
FT sig_peptide 15. 68
FT mat_peptide 69. 968
FT /product="human apolipoprotein E3".
FEATURES
source Location/Qualifiers
1. 1110

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      198 a      355 c      414 g      143 t
ORIGIN

Alignment Scores:
Pred. No.:      1,38e-79      Length:      1110
Score:          1574.00      Matches:      315
Percent Similarity: 99.68%      Conservative: 1
Best Local Similarity: 99.37%      Mismatches: 1
Query Match:      99.31%      Indels:      0
DB:              6      Gaps:      0

US-09-827-854-19 (1-317) x E08423 (1-1110)

QY      1  MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      15  ATGAAGGTTCTGTGGCGCGCTGTGCTACATTCCTGGCAGAGATGCGACGCAAGTG 74
QY      21  GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      75  GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCGACAGACCGAGACTGGCAGAGC 134
QY      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      135  GGCCAGCGCTGGAGACTGGCACTGGGTGCTTTGGGATTACCTGCGCTGGTGCGACACA 194
QY      61  LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      195  CTGCTCAGCAGGTGACAGAGAGAGAGAGCTGCTACGCTCCAGGTCACCCAGAACTAGAGCGC 254
QY      81  LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
      255  CTGATGGACGAGACCATGAAAGAGTGAAGGCCCTACAAATGCGAATCGGAACCAACTG 314
QY      101  ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
      315  ACCCGGTTGGCGGAGAGACCGCGGCTGTCTCAAGGAGGTGCGAGCGCGCGAGGCC 374
QY      121  ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
      375  CGCGTGGCGCGGACATGGAGAGAGTGGCGCGCGCTGCGTGCAGTACCGCGGCGAGAGTG 434
QY      141  GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      435  CAGGCCATGCTCGGCCAGACACCGAGAGACTCGCGGCGCTCGCTCCACCTCGCGC 494
QY      161  LysLeuArgGlnArgLeuLeuArgAspAlaAspLysGlnLysArgLeuAlaValTyr 180
      495  AAGCTGGGTAAAGCGGCTCTCCGAGTGCCTATGACCTGCGAAGCGCGTGGCAGGTAC 554
QY      181  GlnAlaGlyAlaArgGluGlnGlnAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      555  CAGCGCGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCGCATCCGCGAGCGCGCTGGG 614
QY      201  ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
      615  CCCCTGTGTGAACAGGGCGCGGCTCGCGCGCCACTGTGGGCTCCCTGCGCGCGCAACCG 674
QY      221  LeuGlnGluArgAlaGlnAlaThrPGLuArgPheArgAlaArgMetGluGlnMetGly 240
      675  CTACAGAGAGCGGCGCCAGGCTTGGGCGAGCGGCTGCGCGCGGAGAGAGAGATGGGC 734
QY      241  SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
      735  AGCGGAGACCGCGACCGCTGGAGAGGTGAAGAGCAGAGTGGCGCGAGGTGGCGCAAG 794
QY      261  LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
      795  CTGGAGAGAGAGCGCCAGCAGATACGCTTGCAGGCGCAGGCGCTTCCAGGCGCGCTCAAG 854
QY      281  SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
      1110  |||||||

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      198 a      355 c      414 g      143 t
ORIGIN

Alignment Scores:
Pred. No.:      1.64e-79      Length:      1156
Score:          1573.00      Matches:      315
Percent Similarity: 99.68%      Conservative: 1
Best Local Similarity: 99.37%      Mismatches: 1
Query Match:      99.24%      Indels:      0
DB:              6      Gaps:      0

US-09-827-854-19 (1-317) x BD004277 (1-1156)

QY      1  MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      61  ATGAAGTTCCTGTGGCGCTGTGCTACATTCCTGGCAGAGATGCGACGCAAGTG 120
QY      21  GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      121  GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCGACAGACCGAGTGGCAGAGC 180
QY      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      181  GGCCAGCGCTGGAGACTGGCACTGGGTGCTTTGGGATTACCTGCGCTGGTGCGACACA 240
QY      61  LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      241  CTGCTCAGCAGGTGACAGAGAGAGTGTCTACAGCTCCAGGTACCCAGAACTAGAGCGCG 300
QY      81  LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
      301  CTGATGGACGAGACCATGAAAGAGTGAAGGCCCTACAAATCGGAACCTGAGGAACAATG 360
QY      101  ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
      1156  |||||||

FEATURES             source          location/Qualifiers
BASE COUNT      208 a      367 c      432 g      149 t
ORIGIN

Alignment Scores:
Pred. No.:      1.64e-79      Length:      1156
Score:          1573.00      Matches:      315
Percent Similarity: 99.68%      Conservative: 1
Best Local Similarity: 99.37%      Mismatches: 1
Query Match:      99.24%      Indels:      0
DB:              6      Gaps:      0

US-09-827-854-19 (1-317) x BD004277 (1-1156)

QY      1  MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      61  ATGAAGTTCCTGTGGCGCTGTGCTACATTCCTGGCAGAGATGCGACGCAAGTG 120
QY      21  GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      121  GAGCAAGCGGTGGAGACAGACCGGAGCCGAGCTGCCGACAGACCGAGTGGCAGAGC 180
QY      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      181  GGCCAGCGCTGGAGACTGGCACTGGGTGCTTTGGGATTACCTGCGCTGGTGCGACACA 240
QY      61  LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      241  CTGCTCAGCAGGTGACAGAGAGAGTGTCTACAGCTCCAGGTACCCAGAACTAGAGCGCG 300
QY      81  LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
      301  CTGATGGACGAGACCATGAAAGAGTGAAGGCCCTACAAATCGGAACCTGAGGAACAATG 360
QY      101  ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
      1156  |||||||
```

Db 361 ACCCGGTGGCGGAGGACGCGCGGCTGTCCAGAGAGCTGACGGCGCGACGGCC 420
Oy 121 ArtleugluAlaaspMetgluaspValCysglYArgleuValGlnYrArglYgluVal 140
Db 421 CGGCTGGGCGGACATGAGAGAGGTGTGGCGCCCTGGTGCAgTACC GGCGAGGTG 480
Oy 141 GlnAlaMetleugluYglnSerThrGluGluLeuArglValArgleuAlaSerHisLeuArg 160
Db 481 CAGGCATGCTCGGCGCGACAGACGAGAGAGTGGGGTGGCCCTCGGCTCCCACTCGCC 540
Oy 161 LysLeuArgluArgleuLeuArgluValAspAlaaspLeuGlnYrArgleuAlaValTyr 180
Db 541 AAGCTGGTAAAGCGGTCTCTCCGCGATGCGATGACCTGCAGAGAGTCCCTGCGAGTGTAC 600
Oy 181 GlnAlaGluAlaArgluGluGluAlaGluGluYrLeuSerAlaIleArgluGluGluGly 200
Db 601 CAGGCGGGGCGCGCGAGGGCGCGGCGCTCAGCGCCATCCGAGCGCGCTGGG 660
Oy 201 ProleuValgluGlnGluYrValArgAlaAlaThrValGlySerLeuAlaGluPro 220
Db 661 CCCCTGGTGMAACAGGGCGCGGTGCGGCGCCACTGTGGGCTCCCGCGCGCGACCG 720
Oy 221 LeuGlnGluArgluAlaGlnAlaATrpglyGluArgluArgluArgluMetGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGCGCTGGGCGAGCGCTGCGCGCGCGATGAGAGATGGC 780
Oy 241 SerArthrAraspArgluAspGluValYrLeuGluGluValAlaGluValArgAlaLys 260
Db 781 AGCCGACCGCGACCGCTGTGACGAGGTGAAGAGCAGGTGGGAGGTGCGCGCAAG 840
Oy 261 LeuGluGluGlnAlaGlnAlaIleArgluGlnAlaGluAlaIleArgluAlaArgLeuLys 280
Db 841 CTGAGAGAGAGCGCGCCAGCATACGCTGAGCGCGAGCGCTCCAGCGCGCTCAAG 900
Oy 281 SerTrpPheGluProleuValGluAspMetGlnArgluGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTTCGAGCCCTGTGTGAGACATGACGCGCCAGTGGCGCGGTGTGTGAGAAG 960
Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGACAGCTGCGGTGGACACGCGCGCCCTGTGTGCCAGCGCAATTCAC 1011

RESULT 9
AX333278 1157 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 3787 from Patent WO0194629.
DEFINITION AX333278
ACCESSION AX333278
VERSION AX333278.1 GI:18123912
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 3787 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1. 1157
Location/Qualifiers
BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 1,87e-79
Score: 1572.00
Percent Similarity: 99.378
Best Local Similarity: 99.058
Query Match: 99.188

Length: 1157
Matches: 314
Conservative: 1
Mismatches: 2
Indels: 0

DB: 6 Gaps: 0
US-09-827-854-19 (1-317) x AX333278 (1-1157)
Oy 1 MetlysValleuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGTTCGTGTGGCGCTGCGTGTGCTGTGTCATATCTCGCAGAGATCCAGGCCAAGGTG 121
Oy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgluGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAACCGGTGAGACAGAGCCGAGCCGACGCTGCGCCAGCAGACGAGAGTGGCAGAGC 181
Oy 41 GlyGlnArthrpgluLeuAlaLeuGluYrPheThrAspTyrLeuArgluTrpValGlnThr 60
Db 182 GGGCAGCGCTGGAGACTGGCACTGGTCCCTTTGGATTAACCTGCGCTGGGTCCAGACA 241
Oy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGCTGACGAGGCGAGGAGAGAGCTCTCAGCTCCCAAGTCAACCAAGACTGAGGGCG 301
Oy 81 LeuMetAspGluThrMetlysGluLeuLysAlaTyrTrpSerGluLeuGluGluGlnLeu 100
Db 302 CTGATGGACGAGACCATGAAGAGATTGAAGGCTTCAAAATCGAACTGGAGGAACAATCG 361
Oy 101 ThrProValAlaGluGluThrArgluAlaArgLeuSerTrpGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTACCGAGAGAGAGCGCGGCGACGCTGTCCAAAGAGCTGCAGAGCGCGCAGGCC 421
Oy 121 ArgLeuGlyAlaaspMetgluaspValCysglYArgleuValGlnYrArglYgluVal 140
Db 422 CGGCTGGGCGCGACATGAGAGAGAGCTGTGCGCGCGCTGTGTGACAGTACC GGCGAGGTG 481
Oy 141 GlnAlaMetleugluYglnSerThrGluGluLeuArgluValArgleuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGGCGCACAGCACCGAGAGACTGGGGTGGCGCTCGCTCCCACTCGCGC 541
Oy 161 LysLeuArgluArgluLeuArgluAspAlaaspLeuGlnYrAspLeuAlaValTyr 180
Db 542 AAGCTGGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCCCTGGCGAGTTCAC 601
Oy 181 GlnAlaGlyAlaArgluGluGluAlaGluArgGlyLeuSerAlaIleArgluGluGluGly 200
Db 602 CAGGCGGGGCGCGCGAGGCGCGCGAGCGCGCTCAGCGCCATCCCGAGCGCGCTGGG 661
Oy 201 ProleuValgluGlnGluYrValArgAlaAlaThrValGlySerLeuAlaGluPro 220
Db 662 CCCCTGGTGMAACAGGGCGCGGTGCGGCGCCACTGTGGGCTCCCGCGCGCGACCG 721
Oy 221 LeuGlnGluArgluAlaGlnAlaATrpglyGluArgluArgluArgluMetGluMetGly 240
Db 722 CTACAGAGAGCGGGCCAGGCGCTGGGGGAGAGCGGTGCGCGCGGATGAGAGATGGC 781
Oy 241 SerArthrAraspArgluAspGluValYrLeuGluGluValAlaGluValArgAlaLys 260
Db 782 AGTCGAGCCCGCGACCGCTGTGAGAGAGTGAAGAGAGAGTGGCGGAGTGGCGGCCAAG 841
Oy 261 LeuGlnGluGlnAlaGlnAlaIleArgluGlnAlaGluAlaIleArgluAlaArgLeuLys 280
Db 842 CTGAGAGAGAGCGCCAGCATACGCTGACAGCGCGAGCGCTTCCAGGCGCGCTCAAG 901
Oy 281 SerTrpPheGluProleuValGluAspMetGlnArgluGlnTrpAlaGlyLeuValGluLys 300
Db 902 AGCTGTTCGAGCCCTGTGTGAGACATGACAGCGCCAGTGGCGCGGTGTGTGAGAAG 961
Oy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGACAGCTGCGGTGGACACAGCGCGCCCTGTGTGCCAGCGCAATTCAC 1012

RESULT 10
AX409597 1157 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 2244 from Patent WO0229103.
DEFINITION AX409597
ACCESSION

```

VERSION      AX409597.1  GI:21442302
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1
              Alvarres,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
              TITLE      Gene expression profiles in liver cancer
              JOURNAL    Patent: WO 0229103-A 2244 11-APR-2002;
              GENE LOGIC INC (US)
FEATURES
  source      location/Qualifiers
              1..1157
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /note="EMBL/GenBank Accession No. M12529"
BASE COUNT   212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.:    1,87e-79      Length: 1157
Score:        1572.00      Matches: 314
Percent Similarity: 99.37%  Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match:  99.18%      Indels: 0
DB:           Gaps: 0

US-09-827-854-19 (1-317) x AX409597 (1-1157)
QY      1 MetlysvallleutrpalaialeuleuValThrpheleualaglycysglnalalysval 20
        |||||||
Db      62 ATGAAGGTTCTGTGGGCTGTGCTGTGCATTCTCGGAGAGATGCCAGAGCCAGAGGTG 121
QY      21 GluGlnAlaValaGluThrgluProgluProgluLeuAargGlnGlnThrgluTrpGlnSer 40
        |||||||
Db      122 GAGCAAGCGGTGGAGACAGACGCCGAGCTGGCCGACGACGACGACGACGACGACGACGAC 181
QY      41 GlyGlnAargTrpGluLeuAlaLeuGlyAargPheTrpAspTrpLeuAargTrpValGlnThr 60
        |||||||
Db      182 GGCACGCGCTGGCAACCTGGCTGCTTTGGGATTCCTGCGCTGGGTGGAGACA 241
QY      61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrglnGluLeuAargAla 80
        |||||||
Db      242 CTGTCTGACGAGGTGCAGAGAGAGCTGTCACTCCCAAGTACCAGGACCAAGAGTGAAGGCG 301
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
        |||||||
Db      302 CTGATGACGACGACCATGAAGAGTGAAGGCTTACAATCGAAGTGAAGGACCAACTG 361
QY      101 ThrProValAlaGluGluThrAargAlaAargLeuSerLysGluLeuGlnAlaGlnAla 120
        |||||||
Db      362 ACCCGGAGGAGGAGAGACGCGGCGACGCTGTCCAAGAGCTGCAGACGCGCAGAGCC 421
QY      121 AargLeuGlyAlaAspMetGluAspValCysGlyAargLeuValaGlnTrpAargGlyGluVal 140
        |||||||
Db      422 CGGCTGGGCGGGGACATGAGAGAGTGTGGCGCGCTCGGTGACGTACCGGGGAGGTG 481
QY      141 GlnAlaMetLeuGlnGlnSerThrgluGluLeuAargValAargLeuAlaSerHisLeuAarg 160
        |||||||
Db      482 CAGGCAATGCTCGGCGACAGACGAGAGAGTGGGGGTGGCGCTCGCTCCACTGCGC 541
QY      161 LysLeuAargGlnAargLeuLeuAargAspAlaAspAspLeuGlnLysAargLeuAlaValTrp 180
        |||||||
Db      542 AAGCTGGTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGGCTGGCAGTGTAC 601
QY      181 GlnAlaGlyAlaAargGluGluAlaGluAargGlyLeuSerAlaAlaLeuAargGluValTrp 200
        |||||||
Db      602 CAGGCGGGGGCGCGAGGGGCGCGAGCGGCTTCAGCGCATCCGCGAGCGGCTGGGCG 661
QY      201 ProLeuValaGluGlnGlyAargValaAargAlaAargValaGlySerLeuAlaGlyGlnPro 220
        |||||||
Db      662 CCCTTGGTGGAAAGGCGCGGCTGGCGCGCCACTGTGGGCTCCCTGGCGGCGCAGCGC 721

```

```

QY      221 LeuGlnGluAargAlaGlnAlaTrpGlyLeuAargLeuAargMetGluMetLys 240
        |||||||
Db      722 CTACAGAGAGCGGGCCAGGCTGGGGCGAGCGGTGCGCGCGGATGAGAGATGGGC 781
QY      241 SerAargTrpAargAspAargLeuAspGluValLysGluGlnValaGluValaAargAlaLys 260
        |||||||
Db      782 AGTCGACCCCGGACCGCTCTGGACGAGGTGAAGACAGCAGGTGGCGGAGGTGCGGCCAAG 841
QY      261 LeuGluGlnGlnAlaGlnGlnAlaAargLeuGlnAlaAargLeuAlaAargLeuLys 280
        |||||||
Db      842 CTGGAGAGAGAGCGCCACAGATACGCTGCAGGCGGAGGCTTCGACAGCGCGCTCAAG 901
QY      281 SerTrpPheGluProLeuValaGluAspMetLysAargGlnTrpAlaGlyLeuValaGlyLys 300
        |||||||
Db      902 AGCTGTTCGAGCCCTGTGTGTAAGACATGCAGCGCGCAGTGTGGCGCGGCTGTGTGAGAG 961
QY      301 ValGlnAlaAlaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
        |||||||
Db      962 GTGCAGGCTGCGGTGGGACCAAGCGCGCCCTGTGCGCCAGCAATCAC 1012

RESULT 11
LOCUS       115975              1157 bp.  DNA  linear  PAT 03-APR-1996
DEFINITION  Sequence 1 from patent US 5472858.
ACCESSION   115975
VERSION     115975.1  GI:1250883
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE    1 (bases 1 to 1157)
AUTHORS     Attie,A.D., Gretsch,D.G., Sturley,S.L. and Beckage,N.E.
TITLE       Production of recombinant proteins in insect larvae
JOURNAL     Patent: US 5472858-A 1 05-DEC-1995;
FEATURES
  source      location/Qualifiers
              1..1157
              /organism="unknown"
BASE COUNT   212 a 370 c 426 g 149 t
ORIGIN
Alignment Scores:
Pred. No.:    1,87e-79      Length: 1157
Score:        1572.00      Matches: 314
Percent Similarity: 99.37%  Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match:  99.18%      Indels: 0
DB:           Gaps: 0

US-09-827-854-19 (1-317) x 115975 (1-1157)
QY      1 MetlysvallleutrpalaialeuleuValThrpheleualaglycysglnalalysval 20
        |||||||
Db      62 ATGAAGGTTCTGTGGGCTGTGCTGTGCATTCTCGGAGAGATGCCAGAGCCAGAGGTG 121
QY      21 GluGlnAlaValaGluThrgluProgluProgluLeuAargGlnGlnThrgluTrpGlnSer 40
        |||||||
Db      122 GAGCAAGCGGTGGAGACAGACGCCGAGCTGGCCGACGACGACGACGACGACGACGACGAC 181
QY      41 GlyGlnAargTrpGluLeuAlaLeuGlyAargPheTrpAspTrpLeuAargTrpValGlnThr 60
        |||||||
Db      182 GGCACGCGCTGGCAACCTGGCTGCTTTGGGATTCCTGCGCTGGGTGGAGACA 241
QY      61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrglnGluLeuAargAla 80
        |||||||
Db      242 CTGTCTGACGAGGTGCAGAGAGAGCTGTCACTCCCAAGTACCAGGACCAAGAGTGAAGGCG 301
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
        |||||||
Db      302 CTGATGACGACGACCATGAAGAGTGAAGGCTTACAATCGAAGTGAAGGACCAACTG 361
QY      101 ThrProValAlaGluGluThrAargAlaAargLeuSerLysGluLeuGlnAlaGlnAla 120
        |||||||
Db      362 ACCCGGAGGAGAGAGACGCGGCGACGCTGTCCAAGAGCTGCAGACGCGCAGAGCC 421

```


QY	121	ArgLeuGlyAlaAspMetGluAspValYsgGlyArgLeuValGIntyrArgGlyGluVal	140
Db	422	CGGCTGGGGCGGACATGAGAGACTGTGGGCGCCCTGTGCTACATACCGGGCAGAGT	481
QY	141	GlnAlaMetLeuGlyGlnSerThrGlnGlnLeuValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCCATGCTGGGCAACAGACCCGAGGAGAGTGGGGTGGCTGTGCTCCACCTGGGC	541
QY	161	LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	542	AAGCTGCTAAGCGGCTCTCTCCGCAATCCCGCTACCTGCAGAAACGCCCTGGCAGTAC	601
QY	181	GlnAlaGlyAlaArgGlyGlyValAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly	200
Db	602	CAGGCCGGGGCCCGGAGGGCGCCGAGGGCGGCTTCACAGCCCATCCGAGGCGCTGGGG	661
QY	201	ProLeuValGlnGlnGlnIArgValAlaArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	662	CCCCGTGGTGAACAGAGGGCGCCCTGGGGGCGCCCACTGTGGGCTCCCTGGCCGACGCC	721
QY	221	LeuGlnGlnArgAlaGlnAlaThrPylGlnArgLeuArgAlaArgMetGlnLysGly	240
Db	722	CTACAGAGAGCGGGCCCAAGGCTGGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC	781
QY	241	SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys	260
Db	782	AGTCGGACCCCGGACCGCTGTGCAGAGGTGAAGAGCAGGTGGCGAGGTGGCCGCCAAG	841
QY	261	LeuGlnGlnGlnAlaGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys	280
Db	842	CTGGAGGAGCAGAGCCGACAGATAGGCTGTGCAGGCGGAGCCTTCAGAGCCGCTCAG	901
QY	281	SerTrpPheGluProLeuValGlnLysPheGlnArgGlnTrpAlaGlyLeuValGluLys	300
Db	902	AGCTGTTCGAGCCCTGTGTGAACATGCAGGCCAGTGGGCGGGCTGTGTGAGAGAAG	961
QY	301	ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis	317
Db	962	GTGCAGGCTGGCGTGGGCACAGCCGCCCTGTGCTGCACGACAAATCAC	1012
RESULT 12			
LOCUS	HUMAPOE	1157 bp	mRNA linear PRI 08-AUG-1995
DEFINITION	Human apolipoprotein E mRNA, complete cds.		
ACCESSION	M12529		
VERSION	M12529.1	GI:178848	
KEYWORDS	apolipoprotein.		
SOURCE	Homo sapiens (clone: pHAEl112,178,813). male 57-year old liver		
ORGANISM	CDNA to mRNA.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1157)		
JOURNAL	Human apolipoprotein E mRNA. CDNA cloning and nucleotide sequencing		
MEDLINE	J. Biol. Chem. 259 (10), 6498-6504 (1984)		
PubMed	84212473		
FEATURES	6327862		
SOURCE	Location/Qualifiers		
	1..1157		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="19q13.2"		
	/clone="pHAEl112,178,813]."		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="57-year old"		
	1..1157		
	/gene="APOE"		
	<1..1157		
gene			
mrna			

CDs	62..1015	/gene="APOE" /note="apoE mRNA; G00-119-691"
	/gene="APOE" /note="precursor" /codon_start=1 /product="apolipoprotein E" /protein_id="AAB5518.1" /db_xref="GI:178849" /db_xref="GDB:G00-119-691" /translation="MKVLMALVTFLAGCAQKAEVAETEPEBELKQOTEMSGQRM ELAGFEMDYLRMVOTLSEVOEELSSOYQELRALMDETPKELKAVKSELQLP VAEETFRRLSKELQTAORIGADMDQNGRLVYRGVQAMLGSTELRYRLASHLR KLRRRLRPDDLDOKRLAVYQAGAREGERLSIARELGLYVQGRARRATVSLAG QPLERAOAGMERLRLARMEWGSRTRODLDEKQAEVVRKLEQAKQOURLQAEAFQ ARLSWEPELVEDMQROWAGLVEKQVAAGTSAAPVSDNH"	
sig_peptide	62..115	/gene="APOE" /note="G00-119-691"
mat_peptide	116..1012	/gene="APOE" /product="apolipoprotein E" /note="G00-119-691"
variation	370	/gene="APOE" /note="a in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	410	/gene="APOE" /note="g in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	569	/gene="APOE" /note="c in PHAE[112,178]; g in PHAE813; G00-119-691" /replace="g"
variation	784	/gene="APOE" /note="t in PHAE[112,178]; c in PHAE813; G00-119-691" /replace="c"
variation	859	/gene="APOE" /note="g in PHAE[112,178]; a in PHAE813; G00-119-691" /replace="a"
BASE COUNT	212 a 370 c 426 g 149 t	
ORIGIN	30 bp upstream of BstNI site.	
Alignment Scores:		
Pred. No.:	1.87e-79	Length: 1157
Score:	1572.00	Matches: 314
Best Local Similarity:	99.37%	Conservative: 1
Percent Similarity:	99.05%	Mismatches: 2
Query Match:	99.18%	Indels: 0
DB:	9	Gaps: 0
US-09-827-854-19 (1-317) x HUMAPOE (1-1157)		
0Y	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	62 ATGAAGGTTCTGTGGGCTTCGCTGCTGGTACATTCTGGCAGGATGCCAGGCCAAGGTG	121
0Y	21 GluGlnAlaValAlaGluThrGluProGluProGluLeuLeuArgGlnGlnThrGluTrpGlnSer	40
Db	122 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCAGCAGACAGCCAGTGGCAGAGC	181
0Y	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpPaspArgTrpLeuArgTrpValGlnThr	60
Db	182 GCCCAGCCCTGGGAACTGGCACTGGGTGCTTTTGGGATACCTGCGCTGGGTGCAGCA	241
0Y	61 LeuSerGluGlnValAlaGluGluLeuLeuSerSerGlnValAlaThrGlnGluLeuArgAla	80
Db	242 CTGTGTGACAGAGTCCAGAGAGAGCTGCTCAGCTCCCAAGTACCCAGAACTGAGGCGC	301
0Y	81 LeuMetAspGluThrMetLysGluLeuLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu	100

```

|||||
Db 302 CTGATGAGCAGACCATTAAGAGAGTTGAAAGCCCTAACAAATCGGAAGCTGAGAACAACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
Db 362 ACCCGGAGGAGGAGAGAACCGGGGACGGCTGTCCAAAGAGAGCTGCACAGCGCGGAGGCC 421
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 422 CGCGTGGCGGCGACATGAGAGAGCTGTGGCGCCCTGTGGTACAGTACCGGGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTGGCGCAAGACAGCAGAGAGCTGGGGTGGCGCTCCCTCCACCTGGCGC 541
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCGATGACCTGCAGAAAGCGCTGGCACTGAC 601
Qy 181 GlnAlaGlnAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 602 CAGGCGGGGCGCGGAGCGCGCGAGCGGCTCAGCGCATCCGCGAGCGCGCTGGGG 661
Qy 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGGAAACAGGGCGCGTGGGGCCGCCACTGTGGCTCCCTGGCGGCGAGCGG 721
Qy 221 LeuGlnGlnAlaArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGlyGluMetGly 240
Db 722 CTACAGAGAGCGGGGCCAGCGCTGGGGCGAGCGGCTCCCGCGCGGATGGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 782 AGTGGAGACCCCGACCGCGCTGTGAGAGGTGAAGAGAGCGAGTGGGAGAGTGGCGCCAAAG 841
Qy 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGGAGCGAGCGCCACAGATAGCTGTCAGAGCGCGAGCGCTTCACAGCCCGCTCAAG 901
Qy 281 SerTyrPheGluProLeuValGlnAlaAspMetGlnArgGlnTyrAlaGlyLeuValGlnLys 300
Db 902 AGCTGGTTCGAGCGCCCTGTGTGGAAGACATGCAAGCGCCAGTGGCGGCTGTGTGAGAGAG 961
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCAGGCTGCGGTGGGACACAGCGCGCCCTGTGCCACGACGAAATCAC 1012

RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279 Apo E humanized mammal.
DEFINITION BD004279
ACCESSION BD004279.1 GI:18632240
VERSION BD004279.1
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1156)
Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
Apo E humanized mammal
Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PE 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC
A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
C12N15/09//C07K14/775.
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

```

```

FH Key Location/Qualifiers
FEATURES FT CDS Location/Qualifiers
source 1..1156
BASE COUNT 208 a 369 c 432 g 147 t
ORIGIN
Alignment Scores:
Pred. No.: 2,74e-79 Length: 1156
Score: 1569.00 Matches: 315
Percent Similarity: 99.68% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 98.99% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-19 (1-317) x BD004279 (1-1156)
Qy 1 MetLysValLeuTyrPalaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGGCTGGCTGTGCTGATTCATTCCTGGCAGGATGCCAGGCCAAGTGG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrPheGlnSer 40
Db 121 GAGCAACCGGTGAGAGACAGACCCGAGCCGAGCTGGCGCCAGCAGACCGAGTGGCAGAGC 180
Qy 41 GlyLysArgTyrPalaAlaLeuGlnAlaLeuGlnAlaValPheTyrPheTyrPheTyrPheTyrPheTyr 60
Db 181 GGGCAGCGCTGGAGACTGGCACTGGTGGCTTTGGGATTAACCTGGCTGGCTGGCTGGCTGG 240
Qy 61 LeuSerGlnGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTCTGAGCAGGTGCGAGAGAGACTGCTGACAGTCCAGAGCACCCAGAACTGAGGGCG 300
Qy 81 LeuMetSerGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGluLeu 100
Db 301 CTGATGAGCAGACCAATGAGAGAGTTGAAGCCCTCAAAATCGAACTGGAGAGAACTG 360
Qy 101 ThrProValAlaGluGlnGluThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
Db 361 ACCCGGAGGAGAGAGACCGGGGACAGCGCTGTCCAAAGAGCTGCAGCGCGCGGAGGCC 420
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGCGTGGGCGGCGACATGAGAGAGTGGCGGCGCTGTGGTACAGTACCCGCGGAGAGTGG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGCATGCTGGCGCAGAGACAGCAGAGAGCTGGGTGGCGCTCCCTCCACCTGGCGC 540
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGCTCTCCGCGATGCCGATGACTGCAGAAAGCGCCCTGGCAGTAC 600
Qy 181 GlnAlaGlnAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 601 CAGGCGGGGCGCGAGGCGCGAGCGCGAGCGCGCTCAGCGCATCCCGAGCGCCCTGGGG 660
Qy 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValLysSerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGGAAACAGGGCGCGTGGGGCCGCCACTGTGGCTCCCTGGCGGCGGAGCGG 720
Qy 221 LeuGlnGlnAlaArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGGGCCAGCGCTGGGGCGAGCGCTGGCGCGCGGAGAGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
Db 781 AGCGGAGCGCGAGCGCGCTGTGAGAGAGTGAAGAGAGAGTGGCGGAGTGGCGGCGCAAG 840
Qy 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluValPheGlnAlaArgLeuLys 280

```

Db 841 CTGAGAGAGAGAGAGAGAGATACCGCTGACAGCGAGCCCTTCCAGCCCGCCCTCAAG 900
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGGTTCAGAGCCCTGGTGGAGAACATGACAGCCAGAGTGGGGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGACAGCTCCGCTGGGACAGAGCGCCGCTGTGGCCAGCAGCAATATAC 1011
RESULT 14
MFAPOE 1178 bp mRNA linear PRI 31-MAR-1995
LOCUS MFAPOE
DEFINITION Monkey mRNA for apolipoprotein E.
ACCESSION X13887
VERSION X13887.1 GI:38054
KEYWORDS apolipoprotein; apolipoprotein E.
SOURCE Macaca fascicularis.
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Marotti, K.R.
TITLES Direct Submission
JOURNAL Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,
301 Henrietta Street, Kalamazoo, MI 49008
2 (bases 1 to 1178)
AUTHORS Marotti, K.R., Whitted, B.E., Castle, C.K., Polites, H.G. and
Melchior, G.W.
TITLES Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA
JOURNAL Nucleic Acids Res. 17 (4), 1778 (1989)
MEDLINE 89160349
PUBMED 2922300
COMMENT Data kindly reviewed (20-Mar-1989) by Marotti K.R.
FEATURES
source
location/Qualifiers
1..1178
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="1"
/cissue_type="liver"
/clone_lib="lambda gt11"
83..1036
/codon_start=1
/product="proapolipoprotein E (AA 18 to 299)"
/protein_id="CAA32092.1"
/db_xref="GI:38054"
/db_xref="SWISS-PROT:P10517"
/translation="MKVLAALIVTFIACGCAKVRQVPEPEPELRLQAGGSGQPM
EIALGRFMDYLRVQVTLSEVOEELISPVQELTITMDETMKELKAYSELEQISP
VAETRRARLSKELOAAOARLGADMEVRSRLVOYSEVQAMGOSREELRRLASHLR
KLRLRLRDADLOKRLIAYVAGARAGRSALIRERGIPEVQGRVRAAVGSIAS
QPLQERAAQLGERLARKMEKSRTRDRIDEVKAIEVRAKLEQADQISLQAEKFO
ARKLSFEPLVEDMDQWAGLVEKVOAAVASTAPVPIDNH"
sig_peptide
mat_peptide
misc_feature
BASE COUNT 217 a 373 c 426 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-72 Length: 1178
Score: 1447.00 Matches: 295
Percent Similarity: 94.018 Conservative: 3
Best Local Similarity: 93.068 Mismatches: 19
Query Match: 91.298 Indels: 0
DB: 9 Gaps: 0
US-09-827-854-19 (1-317) x MFAPOE (1-1178)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20

Db 83 ATGAGAGTTCCTGGAGCTGGCTGGTGCACATTCCTGGACAGATCCAGGCCAAGGTG 142
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnTrpGlnSer 40
Db 143 GAGCAACCGGTGGAGCCAGAGACGAGACCCGACCTTCGACAGAGCTGAGGCGCAGAGC 202
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspPylrLeuArgTrpValGlnThr 60
Db 203 GGGCAGCCCTGGAGAGCTGGAGCTGGCTGGCTTTTGGATTACTGGCTGGTGGTACACA 262
QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 263 CTGTCTGAGCAGAGTGCAGAGAGAGAGCTGCTACGCCAGAGCTACCCAGAACTGACAGAGC 322
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
Db 323 CTGATGACAGACACATGAGAGAGTGAAGCCCTTACAAATCGAAGCTGAGAGAACAGCTG 382
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 383 AGCCCGGTGGCGAGAGAGAGAGCCGGCGAGCTGTCCAGAGAGCTGCAGGCGCGCAGGCC 442
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 443 CGGCTGGTCCGACATGAGAGAGAGAGCTGCGAGCCGCTGGTGCAGTCCGACAGAGAGTG 502
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 503 CAGGCGATGCTGGGGCAGAGTACCGAGAGCTCGGGCGGCGCTCGCTCCACCTGCGC 562
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 563 AACCTGCGCAGAGCGCTCTCCCGCATGCTGATGACCTGCAGAGAGCCCTGGCAGTGTAT 622
QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 623 CAGGCGGGGCGCGAGAGGCGCGAGAGCGGGGTCAAGCCCATCCGCGAGCGCGCTGGGA 682
QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 683 CCCCTGGTGAAGAGAGGCGCGCTGGCGGGCGCCACCTGTGGCTCCCTGGCGCACCGCG 742
QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 743 CTTCAGAGAGCGGCGCCAGAGCTTGGGTGAGCGCTTCGCCACGATGAGAGATGGGC 802
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 803 AGCGGAGACCGCGACCGCTGGAGAGAGGTGAAGAGAGAGAGTGGCGGAGGTGGCGCCAG 862
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgGluGlnAlaGluAlaPheGlnAlaArgLeuLys 280
Db 863 CTGAGAGAACAGCGCCAGAGATACCTTCAGAGCGAGCGCTTCCAGAGCGCCCTTCAC 922
QY 281 SerTrrPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 923 AGCTGGTTCAGAGCCCTGGTGGAGAACATGACAGCCAGAGTGGGCTGGCTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 983 GTGACAGCTCCGCTGGGACAGAGCGCCGCTGTGGCCATCGACAAATATAC 1033
RESULT 15
AF261279 5491 bp DNA linear PRI 27-OCT-2000
LOCUS AF261279
DEFINITION Homo sapiens apolipoprotein-E gene, complete cds.
ACCESSION AF261279
VERSION AF261279.1 GI:11034800
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Alignment Scores:

Pred. No.:	8,97e-69	Length:	549
Score:	1392.20	Matches:	303
Percent Similarity:	60.80%	Conservative:	1
Best Local Similarity:	60.60%	Mismatches:	2
Query Match:	87.85%	Indels:	194
DB:	9	Gaps:	1

US-09-827-854-19 (1-317) x AF261279 (1-5491)

QY	12	PheuuLaagIcysgInalALySVaIgInuInAlaValGluThrGluProGluProGlu	31
Db	3003	TTCCACACAGAGATGCCAGGCCAAGGTGGAGCAACGCGTGGAGACAGACCGGAGCCGAC	3062
QY	32	LeuArgInglInhrGluTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPhe	51
Db	3063	CTGGCCAGACAGACCGAGTGGCGAGACGGCGACCGCTGGGAACTGGCACTGGGTGGCTTT	31222
QY	52	TrpAspYrTrpLeuArgTrpValGlnThrLeuSerGluGlnValGlnGlnGluLeuSer	71
Db	3123	TGGGATTACTGGCGCTGGGTGGAGACAGTGTGAGACAGTGCAGGAGAGACTGCTGACG	3182
QY	72	SerGlnValThrGlnGluLeu	78
Db	3183	TCCCAAGTCAACCGAGAACTGAGTGAAGTCCCACTCTGGCCTTGACCTCTGTG	3242
QY	78	-----	78
Db	3243	GCGCGCTATACCTCCAGGTCAGGTTTCATTTCGCCCCCTGGCTAAATCTGGGGGG	3302
QY	78	-----	78
Db	3303	CTTGGGTCTGTGCTGAGTTTCACTTCTCCATTTCTGACTCTTG6CTTAACTTC	3362
QY	78	-----	78
Db	3363	TGGAATTCCTCTCTCAGCTTTGTCTCTCTCTCTCCCTTGAGCTCAGTCTCACACT	3422
QY	78	-----	78
Db	3423	CGTCTGGCTGTGTCTGTCTGTCTTCCTCTAGCTCTTTATATAGACAGAGAGATGGGT	3482
QY	78	-----	78
Db	3483	CTCACTGTGTGGCCAGAGCTGGTCTTGAATCTCTGGGCTCAAGCATCTCCCGCTCGG	3542
QY	78	-----	78
Db	3543	CCATCCAAAGTGTGGGATTAGAGCATGAGCACACTTCCGGGCTCTGACTCTTCT	3602
QY	78	-----	78
Db	3603	TCTGTCTGTGCTGTGCCCTCTGTGCATCTGTGCATCTGTCTGTCTCTCTCTCTCG	3662
QY	78	-----	78
Db	3663	GCTCTGTGCCCCGTTCTCTCTCTCTCTCTCTGGGTCTCTGGTCAATCCCATCTGCCCC	3722
QY	78	-----	78
Db	3723	CCCCATCCAGCCCTTCTCCCCGCTCCCATGTGGAGACCCCTCCGCCCTTGCGCG	3782
QY	79	- ArgAlaLeuMetAspGluThrMetGlyGlnLeuGlyAlaTyrLysSerGluLeuGlu	98
Db	3783	CAGGGCCCTATGGACGAGACCACTGGAAGAGTTGAAGGCTTCAATGTGGAACCTGGAGA	3842
QY	98	uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla	118
Db	3843	ACAACGTACCCCGGTGGCGGAGAGACGGGGGACGGCGCTGTCCAAAGAGCTGCAGCGCG	3902
QY	118	aglnAlaArgLeuGlyAlaAspMetLysSpValCysGlyArgLeuValGlnTyrArg	138

[illegible]

Search completed: March 14, 2003, 17:32:49
Job time : 2313.23 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17; Search time 178.728 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585
Sequence: 1 MKVLMAALLVFLAGQAKV.....VEKVOAAGTSAAPVSDNH 317

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=yld
-O=/cgn2.1/USFTO.spool/US09827854.r/unat.11032003_101609_27476/app.query.fasta_1.3576
-DB=N.Geneseq.101002 -QPM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cd1
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09827854.@CGN.1.1.1201.0r/unat.11032003.101609_27476 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MAP -LARGEJUNK -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEEXT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N.Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001C.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1585	100.0	1156	24	AAD22052	Human apolipoprotein E
2	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
3	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
4	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
5	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
6	1581	99.7	1110	7	AAN60409	Human apolipoprotein E
7	1573	99.2	1156	24	AAD22048	Human apolipoprotein E
8	1573	99.2	1156	24	AAD22049	Human apolipoprotein E
9	1573	99.2	1156	24	AAD22051	Human apolipoprotein E
10	1572	99.2	1157	17	AAT06957	Human apolipoprotein E
11	1572	99.2	1157	24	ABN95746	Gene #2244 used to
12	1572	99.2	1157	24	ABN95746	Human benign prot
13	1572	99.2	1157	24	ABN95746	Lung cancer relate
14	1571	99.1	1110	6	AAN50450	Sequence encoding
15	1569	99.0	1156	22	AAR84316	Human ApoE coding
16	1569	99.0	1156	22	AAD22047	Human apolipoprotein
17	1566	98.8	1156	24	AAD22050	Human apolipoprotein
18	1477.5	93.2	1279	22	AAS22437	Human apoE2 gene
19	1458.5	92.0	1107	19	AAX75756	Human apolipoprotein
20	1392.5	87.9	9360	24	ABL31915	Human ApoE genomic
21	1392.5	87.9	10716	24	AAD26034	Human apolipoprotein
22	1387	87.5	3805	20	AAT09524	Human Apo E genomic
23	1387	87.5	3805	20	AAT09526	Human Apo E genomic
24	1365.5	86.2	10716	24	AAD26108	Human apolipoprotein
25	1159.5	73.2	5617	24	AAD32081	Bovine ApoE gene.
26	1159.5	73.2	5617	24	AAD32081	Human alpha-1-anti
27	1159.5	73.2	6026	24	AAD32075	Human albumin prom
28	1134	71.5	1126	19	AAY29159	Nucleotide sequence
29	981	61.9	936	15	AAT09101	ApoA4L2 protease
30	981	61.9	936	17	AAT18070	Human ApoA4L2 CN
31	980	61.8	960	18	AAT18070	Partial human apol
32	943	58.9	597	17	AAT18068	Human ApoA4L2 CN
33	934	58.9	597	15	AAT09099	ApoA4L2 protease CD
34	880	54.7	1381	22	AAS22673	Human cDNA encodin
35	867.5	54.7	786	21	AAT18114	Lung cancer associ
36	674	42.5	600	20	AAT09595	EST clone Cpl47.
37	651	41.1	407	24	ABN95994	Human cDNA for nov
38	625	39.4	478	24	ABN95994	Gene #2492 used to
39	625	39.4	478	24	ABN95994	Colon adenocarcino
40	625	39.4	478	24	ABN95994	Thyroid cancer rel
41	594	37.5	499	22	AAS03049	Human diagnostic a
42	510	32.2	330	12	AAC01980	Human apolipoprote
43	504	31.8	405	21	AAC02139	Human secreted pro
44	485	30.6	345	22	AAN98479	Human EST-derived
45	415	26.2	260	21	AAA40342	Human apolipoprote

ALIGNMENTS

RESULT 1
AAD22052
ID AAD22052 standard; DNA: 1156 BP.

AC AAD22052:
XX 12-FEB-2002 (first entry)

DE Human apolipoprotein E (apoE) allele, apoE2** DNA.

KW Human: apolipoprotein E; apoE: cholesterol; atherosclerosis;
hypertriglyceridaemia; low density lipoprotein; LDL; ds.

OS Homo sapiens.

XX key Location/Qualifiers
FH 61.1014
FT CDS /tag= a

```
FT      sig_peptide      /product= "Human apoe allele, apoe2*"
FT      61..114      /*tag= b
FT      mat_peptide      115..1011
FT      /*tag= c
FT      /product= "Mature human apoe allele, apoe2*"
XX      PN      WO200177136-A1.
XX      PD      18-OCT-2001.
XX      PF      06-APR-2001; 2001WO-US11358.
XX      PR      06-APR-2000; 2000US-0544386.
XX      PR      04-OCT-2000; 2000US-0679088.
XX      PR      05-APR-2001; 2001US-0827854.
XX      PA      (KOSP-) KOS PHARM INC.
XX      PA      (UYBO-) UNIV BOSTON.
XX      PI      Zannis VI, Kyreos KE;
XX      DR      WPI: 2002-010885/01.
XX      DR      P-Psdb: AAE13298.
XX      PT      New apolipoprotein E polypeptide and nucleic acid, useful for lowering
XX      PT      cholesterol, delaying the onset of or treating atherosclerosis in
XX      PT      mammal, without inducing hypertriglyceridemia
XX      PS      Claim 14; Page 83; 91pp: English.
XX      CC      The present sequence is a human apolipoprotein E (apoE)
XX      CC      allele, apoe2* DNA. The apoE lipoproteins are useful for
XX      CC      lowering cholesterol, delaying the onset of atherosclerosis,
XX      CC      treating or regressing atherosclerosis without inducing
XX      CC      hypertriglyceridaemia, in a mammal lacking an endogenous,
XX      CC      normally functioning apoE gene or low density lipoprotein (LDL)
XX      CC      receptor or is at risk for developing atherosclerosis due to
XX      CC      accumulation of lipoprotein remnants in the bloodstream or having
XX      CC      a defect in remnant removal.
XX      SQ      Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other;

Alignment Scores:
Pred. No.: 4,32e-111      Length: 1156
Score: 1585.00      Matches: 317
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 24      Gaps: 0

US-09-827-854-19 (1-317) x AAD22052 (1-1156)
QY      1 MetlysalLeuTrpAlaLeuLeuValThrPheLeuIaGlyCysGlnAlaLysVal 20
DB      61 ATGAAAGGTCTGTGGGCTCTCTGTCACATTCCGGAGATGCCAGCCAGGTG 120
QY      21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnThrGluTrpGlnSer 40
DB      121 GAGCAAGCGGTGAGACAGACGCCGAGCCGAGCTGGCCGACGACGATGGCAGAGC 180
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB      181 GGCCAGCGCTGGAACTGGCACTGGTGGCTTTGGATTACCTGGCTGGTGTCAGACA 240
QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB      241 CTGCTTAGCAGAGTGCAGGAGACCTGCTCAGTCCAGAGTCAACCGAAGTGAAGGCG 300
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB      301 CTGATGACGACGACCATGAAAGAGTTGAAGGCTTACAATAATGGAATGGAGAACTG 360
```

```
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB      361 ACCCGGTGGCGGAGAGACCGGGGACAGGCTGTCCAAAGAGCTGCAGAGCCAGAGCC 420
QY      121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
DB      421 CGGCTGGGCGGACATGAGAGACGTGTGGCGCGCTGGTGTCAATACCGGGGAGAGTG 480
QY      141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB      481 CAGGCCATGCTGGGCCAGAGACAGCAGGAGAGTGGGGTGGCGCTCCGCTCCACCTGCCG 540
QY      161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB      541 AAGCTGGCTCAGCGGCTCTCTCCGATGCCATGACCTGCAGAAAGCCCTGGCAGTGTAC 600
QY      181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB      601 CAGGCGGGGCGCGGAGGCGCGGAGCGCGGCTTCAGCGCATCCGAGCGCGCTGGGG 660
QY      201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB      661 CCCCTGGTGGAAACAGGGCGCGGTGCGGGCCGCCACTGTGGGCTCCGTGGCGGAGAGCG 720
QY      221 LeuGlnGluArgAlaGlnAlaThrGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB      721 CTACAGGAGCGGCGCCAGGCTCGGCGGAGCGGCTGGCGCGCGGAGTGAAGATGGAG 780
QY      241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB      781 AGCCGAGCCCGGACCGCTGTGAGAGGTGAAGAGCAGGAGGAGGAGGAGGAGGAGGAG 840
QY      261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB      841 CTGAGAGGAGCAGGCGCCAGCATATGCTGAGGCGGAGGCGCTTCAGAGCGCGCTCAG 900
QY      281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB      901 AGCTGTGTGAGCCCGCTGTGTGAAGACATGACAGGCCAGTGGCGGCTGTGTGAAGAG 960
QY      301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB      961 GTGCAGGCTGCGTGGGACACAGCGCGCCCTGTGCCACGACCAATCAC 1011

RESULT 2
AAD26035
ID      AAD26035 standard; cDNA; 954 BP.
XX      AC      AAD26035;
XX      DT      26-MAR-2002 (first entry)
XX      DE      Human apolipoprotein E (APOE) cDNA.
XX      KW      Human: anti-lipemic; neuroprotective; noctropic; genetic variant; APOE;
KW      apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
KW      genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
XX      OS      atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
XX      HM      Homo sapiens.
XX      FH      Key
FH      CDS      1..954
FT      /*tag= a
FT      /product= "Human APOE protein"
FT      variation
FT      replace (13, C)
FT      /*tag= a
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      variation
FT      replace (31, G)
FT      /*tag= b
FT      /standard_name= "Single nucleotide polymorphism (SNP)"
FT      variation
FT      replace (364, A)
FT      /*tag= c
```


tumours in an individual via the detection and measurement of the expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179, ABA83181 and ABA83183). The methods of the invention are useful for detecting an ovarian tumour in a patient, for identifying an individual at increased risk for developing ovarian cancer, in prognostic tests for assessing the relative severity of ovarian cancer, in tests for monitoring a patient in remission from ovarian cancer and in tests for monitoring disease status in a patient being treated for ovarian cancer. The methods can additionally be used to identify a particular tumour as being an ovarian tumour (i.e., an epithelial ovarian tumour selected from serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma, mucinous cystadenoma, borderline mucinous tumour, mucinous cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma, clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner tumour. The ovarian tumour marker genes of the invention were identified using SAGE (serial analysis of gene expression) and were found to be overexpressed in a broad variety of ovarian epithelial tumour cells relative to normal ovarian epithelial cells. The marker genes are implicated in immune response pathways, in the regulation of cell proliferation and in protein folding, and many of these are membrane-localised or secreted. In addition to their use as diagnostic and prognostic markers, the ovarian tumour marker genes or their encoded proteins may be used as therapeutic targets for the treatment and prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180, ABA83182 and ABA83184 represent the ovarian tumour marker genes of the invention.

XX Sequence 1147 BP: 210 A; 365 C; 425 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 8.58e-111 Length: 1147
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: Gaps: 0

US-09-827-854-19 (1-317) x ABA83113 (1-1147)

QY 1 MettysValIleuTTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 46 ATGAAAGTTCTGTGGGCTGCTGTCATTCCTCGCAGATGCCAGCAAGGTG 105
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
DB 106 GAGCAAGCGGTGGAGACAGACCCGAGCCGAGCCGACGACGAGAGTGGCAGAGC 165
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyAArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 166 GGGCAGCGCTGGAACTGGCACTGGTGGTGGATTACCTGGCTGGGTGCAGACA 225
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 226 CTGCTGAGCGAGTGCAGAGAGAGCTGCTCAGCTCCACAGTCACCCAGAACTAGGGCG 285
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100
DB 286 CTGATGAGCGAGACCATGAAGAGTTGAAGCCCTACCAATGTGAACTGGAGAACCACTG 345
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 346 ACCCGGCGGCGAGAGACCGGGCGCTGTCAAAGAGCTGCAGCGCGCCAGGCC 405
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 406 CGGCTGGGCGGACATGAGAGACGTGTGCGCGCTGTGTGACGTACCCGCGGAGAGTG 465
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
DB 466 CAGCGCATGTCTCGGCGCAGACGACCGAGAGAGTGGGGGTCCGCTCCCACTGCGC 525
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

|||||:|||||
DB 526 AAGCTGCGTAGCGGCTCTCCGCGATGCCATGACCTGCAGAACCGCTGGAGTGCAC 585
QY 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 586 CAGCGCGGCGCGCGAGGCGCGAGCGCGCTCAGCGCCATCCGAGCGCGCTTGGG 645
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyPro 220
DB 646 CCCCTGGTGGAAACAGGCGCGGTGCGGCGCCACTGTGGGCTCCCTGGCGCCAGCG 705
QY 221 LeuGlnGluArgAlaGlnAlaIleTrpGlyLeuArgLeuArgAlaArgMetGluGluMetGly 240
DB 706 CTACAGGAGCGGCGCGCGAGGCTGGGCGAGCGCTGCGCGCGCGCGATGAGAGATGGGC 765
QY 241 SerArgTrpArgAspArgValAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 766 AGCGGAGCGCGCGCGCGCTGGAGAGGTGAAGAGCAGGTGGGAGGTGGCGCCAG 825
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 826 CTGAGAGAGAGCGCGCGCAGACATACCTGCAGCGCCAGCGCTTCCAGCGCCGCTCAG 885
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 886 AGCTGTTGCGAGCCCGTGGTGAAGACATGCAGCGCCAGTGGCGCGCGGTGGAGAG 945
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 946 GTGCAGGCTGCCGTGGGACAGCGCGCCCTGTGGCCAGCGCAATAC 996

RESULT 5

AAF84315
ID AAF84315 standard; cDNA; 1156 BP.

XX AAF84315;

DT 21-JUN-2001 (first entry)

DE Human Apoe3 coding sequence.

KW Human; Apoe3; Alzheimer's disease; arteriosclerosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 61..1014

FT /*tag= a

FT /product= "Human Apoe3"

PN JP2001017028-A.

XX 23-JAN-2001.

PD 28-APR-2000; 2000JP-0128919.

PF 06-MAY-1999; 99JP-0125647.

PR (MITU) MITSUBISHI CHEM CORP.

PA WPI; 2001-285406/30.

XX P-PSDB; AAB80997.

DR New apoe humanized mammalian cell useful for screening for agents

PT useful for treating or preventing Alzheimer's disease and

PT arteriosclerosis -

XX Disclosure; Page 13-14; 22pp; Japanese.

PS The present invention relates to an Apoe humanised mammalian cell. The

CC present sequence is the coding sequence for human Apoe3, which was used

CC in the method of the present invention. The Apoe humanised mammalian cell

CC can be used for screening for agents useful for treating or preventing

CC Alzheimer's disease and arteriosclerosis.

XX Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.: 8.65e-111 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 22 Gaps: 0

US-09-827-854-19 (1-317) x AAF84315 (1-1156)

QY 1 MetlyValLeuTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGTGGCTGTGTCACATTCCGGCAGATGTCAGGCCCAAGTGTG 120
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCCAGACAGACCAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspTyrLeuArgTrpAlaGlnThr 60
Db 181 GGCAGCGCTGGAGACTGGCTGGCTTTGGGATTAACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTGTGACAGGTGACAGAGAGAGTGTCACTCCAGTCCAGCCAGAACACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGGACGAGACCATGAGAGAGTTGAAGGCTTAAATCGGAACCTGAGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGGAGAGACCGGGGACGCGTGTCCAAGAGACTGACAGCGCGCAGGCC 420
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlnArgLeuValGlnTyrArgGlnVal 140
Db 421 CGCGTGGCGGCGACATGAGAGAGCTGTGGCGCCCTGTGGTACACCGCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGACAGCAGCAGGAGCGTGGCGCTCGCTCCACCTCGCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValLys 180
Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGTAC 600
QY 181 GlnAlaGlnAlaArgGluGlnAlaGlnArgLysLeuSerAlaLeuArgGlnArgGln 200
Db 601 CAGCGCGGGCGCGGAGGCGCGGCGGCGCTCACGCGCATCCGCGAGCGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGGGAAACAGGGCGCGTGGGGCCGCCACTGTGGGCTCTCGTGGCGGCAAGCG 720
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGGAGCGGGCGCGGCGCTGGGGCGAGCGGCGCGCGGCGGATGGAGAGATGGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 781 AGCCGGAGCCCGGACCGCTGTGAGAGGTGAAGAGAGGTGGCGGAGGTGCGGCCCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAGAGGAGCGAGGCCAGCAGATACGCTGTGAGCGCGGAGGCTTCCAGAGCCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGTTCGAGCCCTGTGGTGAAGACATGACGCGCCAGTGGCGGCGGCTGTGTGAGAAAG 960

QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGCGCTGGCGTGGCGACAGCGCCCGCTGTGGCCAGCAGAACATCAC 1011

RESULT 6

AAD22048 standard; DNA; 1156 BP.

AAD22048:

12-FEB-2002 (first entry)

Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.

Human: apolipoprotein E; apoE: cholesterol; atherosclerosis;

hypertriglyceridaemia; low density lipoprotein; LDL; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 61..1014 /tag= a /product= "Human apoE isoprotein, apoE3"

FT sig_peptide 61..114 /tag= b

FT mat_peptide 115..1011 /tag= c

FT /product= "Mature human apoE isoprotein, apoE3"

XX W0200177136-A1.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-US11358.

XX 06-APR-2000; 2000US-0544386.

XX 04-OCT-2000; 2000US-0679088.

XX 05-APR-2001; 2001US-0827854.

XX (KOSP-) KOS PHARM INC.

XX (UTBO-) UNIV BOSTON.

XX Zannis VI, Kypros KE;

XX WPI; 2002-010885/01.

XX P-PSDB: AAE13294.

XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering

XX cholesterol, delaying the onset of or treating atherosclerosis in

XX mammal, without inducing hypertriglyceridaemia

XX Claim 14: Page 81; 91pp; English.

XX The present sequence is a human apolipoprotein E (apoE)

XX isoprotein, apoE3 DNA. The apoE lipoproteins are useful for

XX lowering cholesterol, delaying the onset of atherosclerosis,

XX treating or regressing atherosclerosis without inducing

XX hypertriglyceridaemia, in a mammal lacking an endogenous,

XX normally functioning apoE gene or low density lipoprotein (LDL)

XX receptor or is at risk for developing atherosclerosis due to

XX accumulation of lipoprotein remnants in the bloodstream or having

XX a defect in remnant removal.

SO Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;

Alignment Scores:

Pred. No.: 8.65e-111 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 24 Gaps: 0

Qy	121	ArgLeuGluYAlaAspMetGlnAspVtCysGlyIArgLeuValGlnIYrArgGlyVal	140
Db	421	CGCGTGGCGCGGACATGAGAGACGTGTGCGCGCCCTGTGTGCAATCCGGCGAGCTG	480
Qy	141	GlnAlaMetLeuGlnSerThrGluGluDeuAryValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATGCTCGGCGACAGACCGAGAGAGCTGGCGGGGCGCTCGCTCCACCTCGC	540
Qy	161	LysLeuAArgGlnAArgLeuLeuAryAspAlaAspAspLeuGlnIYrSAArgLeuAlaValTyr	180
Db	541	AAGCGCTTAAGCGGCTCTCCCGGATCCCATATCCTGCAGAGTCCCTGGCAGTGTAC	600
Qy	181	GlnAlaGlyAlaArgGluGluYAlaGluAryGlyLeuSerAlaIleArgIValArgLeuGly	200
Db	601	CAGCGCGGGGCGCGGAGGGCGCGGAGCGCGCTCAACCGCATCCGACCGCCTTGGG	660
Qy	201	ProLeuValGluGlnGlyAryValAryAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCCGTGTGAACAGGGCGCGGTGCGGGCCGCACTGCGGCTCCCTGGCGGCGACCG	720
Qy	221	LeuGlnGluAryAlaGlnAlaIleTrrGlyGluAryGluAryAlaAArgMetGluGluMetGly	240
Db	721	CTACAGGAGCGGGCGCCAGCGCTGGGGCGAGGGCTGGCGCGCGCGAGAGAGATGGC	780
Qy	241	SerArgThrArgAspArgLeuAspGluValIYrGluGlnGluAlaIleValAryAlaIYs	260
Db	781	AGCCGAGACCCCGACCGCTGTGACAGAGGTGAGAGCAAGTGGCGGAGTGGCGCCAC	840
Qy	261	LeuGluGluGlnAlaGlnGlnIleAryGluGlnAlaGluAlaPheGlnAlaAryLeuIYs	280
Db	841	CTGGAGGAGACGCGCCACAGATAGCTGCAGGCGCGAGGCTTCAGAGCGCGGCTCAG	900
Qy	281	SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIYs	300
Db	901	AGCTGGTTCGAGCCCTGTGGTGAAGACATGCAGCGCCAGTGGCGCGGTGTGTGAGAG	960
Qy	301	ValGlnAlaIleValGlyThrSerAlaAlaProValProSerAspAsnHis317	
Db	961	GTCGAGGCTGCGCGTGTGGCACCAGCGCGCGCTGTGCTGCCACGACATCTCAC1011	
RESULT	8		
AAD2049			
ID	AAD2049	standard; DNA; 1156 BP.	
XX	AAD2049;		
AC			
XX	12-FEB-2002	(first entry)	
DT			
XX			
DE	Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.		
XX			
KW	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;		
KW	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
XX			
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	61..1014	
FT		/*tag= a	
FT	sig_peptide	/product= "Human apoE isoprotein, apoE2"	
FT		61..114	
FT	mat_peptide	/*tag= b	
FT		115..1011	
FT		/*tag= c	
FT		/product= "Mature human apoE isoprotein, apoE2"	
XX			
XX	MO200177136-A1.		
PN			
PD	18-OCT-2001.		
XX			
PF	06-APR-2001; 2001WO-US11358.		
XX			
PR	06-APR-2000; 2000US-054386.		

PR	04-OCT-2000; 2000US-0679088.
PR	05-APR-2001; 2001US-0827854.
XX	(KOSP-) KOS PHARM INC.
PA	(UYB-) UNIV BOSTON.
PI	Zannis VI, Kypros KE;
XX	WPI: 2002-010885/01.
DR	P-PSDB; AAEI3295.
XX	
PT	New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT	cholesterol, delaying the onset of or treating atherosclerosis in
PT	mammal, without inducing hypertriglyceridemia -
PS	Claim 14; Page 81-82; 91pp; English.
XX	
CC	The present sequence is a human apolipoprotein E (apoE)
CC	isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC	lowering cholesterol, delaying the onset of atherosclerosis,
CC	treating or regressing atherosclerosis without inducing
CC	hypertriglyceridemia. In a mammal lacking an endogenous,
CC	normally functioning apoE gene or low density lipoprotein (LDL)
CC	receptor or is at risk for developing atherosclerosis due to
CC	accumulation of lipoprotein remnants in the bloodstream or having
CC	a defect in remnant removal.
XX	
SQ	Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other:
Alignment Scores:	
Pred. NO.:	3,48e-110 Length: 1156
Score:	1573.00 Matches: 315
Percent Similarity:	99.68% Conservative: 1
Best Local Similarity:	99.37% Mismatches: 1
Query Match:	99.24% Indels: 0
DB:	24 Gaps: 0
US-09-827-854-19 (1-317) x AMD22049 (1-1156)	
OY	1 MetLysValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaLysVal 20
Dd	61 ATGAAGTTCTGTGGGCTGGTGGTGTCACATTCTCGGAGATGCCAGGCCAAAGTG 120
OY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Dd	121 GAGCAACGGGTGGAGACAGACCAGCCGCAGACTCCGCACACAACCGAGTGCAGAGC 180
OY	41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Dd	181 GGCCAGCGGTGGAACTGGCACGCGGGTCCCTTTGGATTACTTCGCTGGGTGCAGACA 240
OY	61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgLa 80
Dd	241 CTGTCTGAGCAGAGTGCAGAGAGAGAGCTCTCAGCTCCCAAGTAGTACCAGAGACTAGGGCG 300
OY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTryLysSerGluLeuGluGluGlnLeu 100
Dd	301 CTGATGAGCAGAACCATGAAGGAGTTGAAGGCTTCAATAATCGAATCGAGAGAACCACTG 360
OY	101 ThrProValAlaGluGluThrArgAlaAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Dd	361 ACCCGGTGGCGGAGAGAGACGCGGGCAGCGCTGTCCAAGACACTCTAGCGCGGCAGGCC 420
OY	121 ArgLeuGlnValAspMetGluAspValLysGlyArgGluValGlnTrpArgGlyGluVal 140
Dd	421 CGGTGGGGCCGGACATGGAGGACGTTGCGGGCGCTGTGTCACTATCCGGCGGCAAGTGTG 480
OY	141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Dd	481 CAGGCCATGCTGGGCCAGAGACACCGAGAGAGTGGGGTGGCTGTGCCACACTGGCC 540
OY	161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180

|||||
DB 721 CTACAGAGCGGCCCGAGGCTGGGGGAGCGGCTGGCGCGCGGAGTGAAGATGGGC 780
QY 241 SerATgThrArgAspArgLeuAspGluValIysGluGlnValAlaGluValArgAlaIys 260
DB 781 AGCGGAGCCCGGACCGCCCTGGAGAGAGTGAAGAGAGAGTGGGGAGTGGCGGCAAG 840
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleGlnAlaArgLeuIys 280
DB 841 CTGGAGAGAGAGCGCCAGCATACGGCTTCAGGCCGAGGCTTCAGGCCGCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValIleGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
DB 901 AGCTGTTCCAGGCCCTGGTGGAAAGATCAGCGCCAGTGGCGCGGCTGGTGGAGAG 960
QY 301 ValGlnAlaIleValIleGlyThrSerAlaIleProValProSerAspAsnHis 317
DB 961 GTGCAAGCTGCGCTGGGACACAGCGCCGCTGTGCCAGCAATTCAC 1011
RESULT 10
AAT06957
ID AAT06957 standard: cDNA to mRNA; 1157 BP.
XX AAT06957;
AC 19-JUN-1996 (first entry)
XX
DE Human apolipoprotein-E (ApoE) cDNA.
XX
KW Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;
KW Manduca sexta; Autographa californica nuclear polyhedrosis virus;
KW haemolymph; lipid complex; biologically active; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc_feature 62..119
FT note="tag= a
FT note="misc_signal"
FT CDS 62..1015
FT /tag= b
XX
XX US5472858-A.
XX
XX 05-DEC-1995.
XX
XX 04-JUN-1991: 91US-0709949.
XX
XX 04-JUN-1991: 91US-0709949.
XX
XX (MISC) WISCONSIN ALUMNI RES FOUND.
XX
XX Attie AD, Beckage NE, Gretch DG, Sturley SL;
PI
XX MPI, 1996-029812/03.
DR P-PSDB; AAR86791.
XX
XX Prod. of recombinant apo:lipoprotein E in insects - by infecting
PT Manduca sexta larvae with recombinant Autographa californica
PT nuclear polyhedrosis baculovirus vector.
XX
XX Disclosure; Columns 11-14; 10pp; English.
XX
XX Recombinant human apolipoprotein-E (ApoE) (AAR86791) can be produced
CC by preparing a genetic construct (contg. an ApoE-encoding sequence,
CC e.g. AAT06957, and flanking regulatory sequences enabling the protein
CC to be expressed in insect cells), which is then introduced into a
CC Manduca sexta larva (using a recombinant Autographa californica
CC nuclear polyhedrosis virus) and recovering the protein from the
CC haemolymph of the larval host. The ApoE produced is in a form
CC sufficiently complexed with lipids to be biologically active, which
CC cannot be achieved in insect cell cultures, and can therefore be
CC used in therapeutic applications.

XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
Alignment Scores:
Pred. No.: 4,15e-110 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 17 Gaps: 0
US-09-827-854-19 (1-317) x AAT06957 (1-1157)
QY 1 MetIysValLeuTrpAlaIleAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
DB 62 ATGAGGTTCTGGGGCTGGCTGGCTGTCTGTCATCTCTCGAGAGATGCCAGGCCAAGGTG 121
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAACCGGTGGAGAGACAGACCGGAGCCGAGCTGCGCCAGAGACCGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GGCAGAGGCTGGGACACTGGCAGCTGGCTTTGGGATTACTGCCCTGGTGGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGCTCAGCAGAGTGCAGAGAGCTGCTCAGCTCCCAATCCCAAGTCAAGAGGGCG 301
QY 81 LeuMetAspGluThrMetIysGluLeuValAlaThrIysSerGluLeuGluGlnGlnLeu 100
DB 302 CTGATGACAGACAGCAATGAAGAGTTGAAGCCCTACAAATCGAATCGAGGAGCAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaIleGlnAla 120
DB 362 ACCCGGTTAGCGAGAGAGACCGCGCAGCGCTTCCAGAGAGCTGCAGACGGCGAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 422 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCGCGCTGTGCAGTACC CGCGGAGGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTCGGCCAGACAGACAGAGAGTCTGGTGGCTTCCTCCGCTCCACCTGGCC 541
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnIysArgLeuAlaValTrp 180
DB 542 AACCTGGTAAAGCGGCTCCCTCCGCAATCCGATGACCTCGAGAAGCGCTGGCAGTATC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGCGAGCGCGCGAGCGCGGCTCAGCGCCATCCGCGAGCCCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 662 CCCCTGGTGAACAGAGGCCCGCTGGCGGCGCCCACTGTGGGCTCCCTGGCCGCGCAGCC 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 722 CTACAGAGAGCGGCCAGAGCTGGGGAGAGCGGCTGGCGCGCGGATGAGAGATGGGC 781
QY 241 SerATgThrArgAspArgLeuAspGluValIysGluGlnValAlaGluValArgAlaIys 260
DB 782 AGTCGACCGCGCAGCGCTCGAGAGTGAAGAGAGAGTGGCGGAGTGGCGCGCCCAAG 841
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleGlnAlaArgLeuIys 280
DB 842 CTGGAGAGAGAGCGCCAGCATACGGCTTCAGGCCGAGGCTTCAGGCCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValIleGluAspMetGlnArgGlnTrpAlaGlyLeuValGluIys 300
DB 902 AGCTGTTCCAGGCCCTGGTGGAAAGATCAGCGCCAGTGGCGCGGCTGGTGGAGAG 961


```
OY 301 ValGlnAlaIaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGACAGGCTGCGGTGGGACACCGCCGCCCTGTGTGCCAGGACATCACC 1012

RESULT 11
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2244 used to diagnose liver cancer.
XX
KW Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
DR WPI, 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX
PS Claim 1; SEQ ID NO 2244; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SO Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 4,15e-110 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
Gaps: 0
DB: 24

US-09-827-854-19 (1-317) x ABN95746 (1-1157)

OY 1 MelysValLeuTrpAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGGCTGTGCTGTGCATTCTGTGGCGAGATGCGACGCCAAGGTG 121

OY 21 GlnGlnAlaIaValaGlyThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGACAGCCGAGCGAGCTGCGCAGACAGACGAGTGGCAGAC 181
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GCGCAGCGCTGGAACTGTGCACCTGGTCTTTGGATTACTGCTGCGTGGTGACAGA 241
OY 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
DB 242 CTGTGTGACAGGTGTGAGAGAGAGCTGCTACACTCCCAATACCCAAACACTGAGAGGCG 301
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyLysSerGluLeuGlnGluGlnLeu 100
DB 302 CTGATGAGACGACATCAGATGAGAGAGTGAAGGCTTACAAATTCGAACTGGAGACACTG 361
OY 101 ThrProValaIaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 362 ACCCGGTAGCGAGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGACGCGCAGGCGC 421
OY 121 ArgLeuGlyAlaAspMetGluLysValaCysGlyArgLeuValaGlnThrArgGlyGluVal 140
DB 422 CGGCTGGGCGCGACATGAGAGAGCTGTGCGGCGCGCTGTGTACAGTACCGCGGAGGTG 481
OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValaArgAlaSerHisLeuArg 160
DB 482 CAGGCCATCTCTGCGCAGACAGCAGCGAGCTGCGGCTGCGCTCCCTCCACCTGCGGC 541
OY 161 LysLeuArgGlnArgLeuLeuArgAlaAspAspLeuGlnLysArgLeuAlaValaTy 180
DB 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTCCAGAAAGCGCTGCGACAGTAC 601
OY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGCAGCGCGCTCAGCGCATCCGCAAGGCGCGTGGGG 661
OY 201 ProLeuValaGluGlnGlyArgValaArgAlaIaThrValaGlySerLeuAlaGlyGlnPro 220
DB 662 CCCCTGTGTGAAACAGAGCGCGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
OY 221 LeuGlnGluArgAlaGlnAlaIaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 722 CTACAGAGAGCGGCGCAGGCTGTGGGCGAGCGGTGCGCGCGATGAGAGAGATGGGC 781
OY 241 SerArgThrArgAspArgLeuAspGluValaLysGluGlnValaIaGluValaArgAlaLys 260
DB 782 ACTCGGACCGCGGACCGCTGTGACAGAGGAGGAGCAAGTGGCGAGCTGCGCGCCAG 841
OY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGAGACAGCGCCAGCAGATACGCTGCGAGCGCGAGGCTTCCAGGCGCCGCTCAG 901
OY 281 SerTrpPheGluProLeuValaGluLysPheGlnArgGlnIleTrpAlaGlyLeuValaGlyLys 300
DB 902 AGCTGTTGTGACCGCTGTGTGAACAATGACAGCGCAATGGCGGCTGTGTGAGAAAG 961
OY 301 ValGlnAlaIaValaGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGACAGGCTGCGGTGGGACACCGCCGCCCTGTGTGCCAGGACATCACC 1012

RESULT 12
ABK64514
ID ABK64514 standard; DNA; 1157 BP.
XX
AC ABK64514;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human benign prostatic hyperplasia gene #409.
XX
KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS Homo sapiens.
```

XX WO200212440-A2.
PN 14-FEB-2002.
XX
XX 07-AUG-2001; 2001WO-US24708.
XX
XX 07-AUG-2000; 2000US-22323P.
PR 05-JUN-2001; 2001US-0873319.
XX
XX (GENE-) GENE LOGIC INC.
PA (NISB) JAPAN TOBACCO INC.
XX
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
DR WPI: 2002-257476/30.
XX
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
XX
XX
PS Disclosure: Page 239-240; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:

Alignment Scores:
Pred. No.: 4,15e-110 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-19 (1-317) x ABK64514 (1-1157)
QY 1 MetLysValLeuThrPAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGGGCTCGTGGTGCACATTCCTGCGAGGATCCAGGCCAAGGTG 121
QY 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACACAGCCGAGCCGAGCTCGGCAGACAGCCAGTGGCAGAC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCACAGCGCTGGGAACCTGCGACTGCGCTTTGGATTACCTGCGCTGGGCGAGACA 241
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||

DB 242 CTGTCTGACAGGTGACGAGAGAGCTGCTACGCTCCCAAGTCAACCCAGAACTGAGGGC 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB 302 CTGATGGAGAGACCATGAAAGGCTTGAAGGCTTACAAATTCGGAATCGAGAGAACATG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCCGCTAGCGAGAGACACCGGCGACGCTCTCCAAAGAGCTGACAGCGGCGAGCC 421
QY 121 ArgLeuGluValAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
DB 422 CGGCTGGCGCGACATGAGAGACGTGTGCGCGCTGCTGTCAGTACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGTGGCCAGAGACCCAGAGAGCTGGGGTGGCTGCTGCCATCCAGCTGGCC 541
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
DB 542 AAGCTCGTAAAGCGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGluValAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGCGCGGGCCCGCGAGGCGCGAGCGCGCTCAGCGCCATCCGAGCGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCGTGGCGCGCCACATGTGGGCTCCCTGGCGCGCGCG 721
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnGly 240
DB 722 CTACAGAGACGGGGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGGTGAGAGAGAGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCGAGACCCGCGAGCGCTGAGAGAGGTGAAGAGACAGTGGCGAGGTGCCGCGCAAG 841
DB 842 CTGAGAGACAGGCGCCAGCAGATACGCTGAGCGAGCGAGCGCTTCCAGCGCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
DB 902 AGCTGTTGTGAGGCTCTGTGTGAAGACATGCAAGCCGAGTGGCGCGGTGTGGAGAG 961
QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProAlProSerAspAsnHis 317
DB 962 GTGCAGGCTGCGGTGGGCAACAGCGCGCTGCTGCTCCAGCAGACATCAC 1012

RESULT 13
ABL65450
ID ABL65450 standard; DNA; 1157 BP.
XX
AC ABL65450;
XX
XX 15-MAY-2002 (first entry)
DT
DT Lung cancer related gene sequence SEQ ID NO:3787.
XX
XX
DE
DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
XX Homo sapiens.
OS
PN WO200194629-A2.
XX
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX

PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Sopet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 XX Claim 1; SEQ ID 3787; 44pp; English.
 XX
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
 Alignment Scores:
 Pred. No.: 4.15e-110 Length: 1157
 Score: 1572.00 Matches: 314
 Percent Similarity: 99.37% Conservative: 1
 Best Local Similarity: 99.05% Mismatches: 2
 Query Match: 99.18% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-19 (1-317) x ABL65450 (1-1157)
 QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
 DB 62 ATGAAGGTTCTGTGGGCTGCGTGTGGTGCACATTCCTCGGAGATGCCAGGCCAAGG 121
 QY 21 GtGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGTGCAGACAGACGAGCGAGAGCA 181
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 182 GGCAGCGCTGGGAACCTGGCACTGGGTGCTTTGGATTACTCGCTGGGTGCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 242 CTGTGTGACAGAGTCCAGAGAGAGCTGCTCAGCTCCCAAGTACCAGACTGAGAGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 DB 302 CTGATGAGACAGACCATGAGAGAGTTGAAGGCTCAATCGAATCGAAGTGGAGAACACTG 361
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 362 ACCCGGTAGCGAGAGAACGCGGCGACGCTGCTCAAGAGACTCAGACGCGCAGAGCC 421
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyVal 140
 DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGGCGGCTGTGTGCAATCCAGGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 482 CAGGCCATGCTGGCGCAGAGACCCAGAGAGTGGCGGTGCGGCTGCCACCTGGCGC 541
 QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGCACTGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGlnValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 602 CAGCGCGGGGCGCGGAGGCGCCGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 DB 662 CCCTGTGTGAACAGAGCGCGCGTGGCGCGCCACTGTAGGGCTCCCTGCGCGGCGAGCCG 721
 QY 221 LeuGlnGluArgAlaGlnAlaThrProGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
 DB 722 CTACAGGAGCGGCGCCAGGCGCTGGGGCAGCGCGTGGCGCGCGGATGAGAGATGGGCG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
 DB 782 AGTCGAGACCCGCGACCGCTGAGAGAGTGAAGAGACAGAGTGGCGAGTGGCGCCCAAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 842 CTGAGGAGAGAGCCACAGATGAGTACGCTGAGCGCGGAGGCTTCAGAGCCGCTCAAG 901
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 902 AGCTGTTTCGAGGCCCTGTGTGGAACATGACGCGCCAGTGGGCGGCTGTGTGAGAG 961

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1885

Sequence: 1 MKVLMALVFLAGCAKV.....VEKVOAVGTSAPVPSDNH 317

Scoring table:

	BLAST62
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/runat_11032003_101611_27510/app_query.fasta.1.3576
-DB=issued_patents.NA -QMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USER=US09827854.qcgn.1.1.2254 &runat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO.XLPPY -NO.MMAP -LARGEOUERT -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1572	99.2	1157	1	US-07-709-949-1
2	1134	71.5	1126	4	US-08-949-155-5
3	1134	71.5	1126	4	US-09-819-964-5
4	998	63.0	4267	4	US-08-949-155-51
5	998	63.0	4267	4	US-08-949-155-51
6	980	61.8	660	2	US-08-726-306A-28
7	507	32.0	330	1	US-07-849-389-6
8	370	23.3	252	3	US-08-617-256-24
9	370	23.3	252	4	US-09-287-141-24
10	370	23.3	252	4	US-09-431-613-24
11	370	23.3	252	4	US-09-504-245-24
12	370	23.3	252	4	US-09-287-682-24

13	370	23.3	252	4	US-09-287-679-24	Sequence 24, Appl
14	370	23.3	252	4	US-09-397-766-24	Sequence 24, Appl
15	370	23.3	252	4	US-09-287-681-24	Sequence 24, Appl
16	370	23.3	252	4	US-09-495-444-24	Sequence 24, Appl
17	176.5	11.1	842	1	US-08-952-796-1	Sequence 1, Appl
18	173.5	10.9	801	4	US-07-959-946-4	Sequence 4, Appl
19	173.5	10.9	801	1	US-08-333-577-4	Sequence 4, Appl
20	173.5	10.9	801	5	PCT-US93-08634-4	Sequence 4, Appl
21	159	10.0	5661	4	US-08-938-105-2	Sequence 2, Appl
22	156.5	9.9	964	4	US-08-448-606-5	Sequence 3, Appl
23	155	9.8	3256	2	US-08-968-751-3	Sequence 3, Appl
24	154	9.7	1879	4	US-09-750-580-2	Sequence 2, Appl
25	153.5	9.7	863	1	US-08-448-606-7	Sequence 7, Appl
26	149	9.4	6306	1	US-08-195-487-3	Sequence 3, Appl
27	149	9.4	6306	5	PCT-US93-06160-3	Sequence 3, Appl
28	146.5	9.2	603	4	US-08-952-726-14	Sequence 14, Appl
29	145	9.1	516	4	US-09-183-861-34	Sequence 34, Appl
30	145	9.1	516	4	US-09-022-765-34	Sequence 3, Appl
31	145	9.1	6306	1	US-08-466-390-3	Sequence 3, Appl
32	145	9.1	6306	1	US-08-470-950-3	Sequence 3, Appl
33	145	9.1	6306	1	US-08-467-781-3	Sequence 3, Appl
34	145	9.1	6306	2	US-08-483-924-3	Sequence 3, Appl
35	144	9.1	1771	2	US-08-533-669A-7	Sequence 7, Appl
36	144	9.1	1771	2	US-08-511-872-1	Sequence 1, Appl
37	144	9.1	1771	4	US-09-183-861-7	Sequence 7, Appl
38	144	9.1	1771	4	US-09-023-765-7	Sequence 7, Appl
39	144	9.1	1960	2	US-08-533-306A-1	Sequence 1, Appl
40	144	9.1	1960	2	US-08-742-923A-1	Sequence 1, Appl
41	144	9.1	8789	1	US-08-328-254-5	Sequence 5, Appl
42	144	9.1	10136	1	US-08-353-700-2	Sequence 2, Appl
43	144	9.1	10136	5	PCT-US95-16216-2	Sequence 126, App
44	144	9.1	13121	4	US-08-961-527-126	Sequence 1, Appl
45	143.5	9.1	11236	1	US-07-853-913-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1157 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 62..1015
: NAME/KEY: misc.signal
: LOCATION: 62..119
: PUBLICATION INFORMATION:
: AUTHORS: McLean, J W
: JOURNAL: Journal of Biological Chemistry
: VOLUME: 259
: PAGES: 6498-6504
: DATE: 1984
: RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
: US-07-709-949-1

```

Alignment Scores:

```

Pred. No.: 1,45e-146 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 1 Gaps: 0

```

US-09-827-854-19 (1-317) x US-07-709-949-1 (1-1157)

```

Qy 1 MetLysValleuTrpAlaIaleuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
Db 62 ATGAAAGTTCTGTGGCTGCTGTGTGTCACATTCCTGCGAGAGTCCAGCGCAAGGTG 121
Qy 21 GUGGAlaValAGluThrGluProGluProGluLeuArgGlnGlnThrGluProGlnIns 40
    |||||||
Db 122 GAGCAAGCGGTGAGACAGACGCGGAGCCGAGCTCGCGCAGACAGCCAGTGGCAGAGC 181
Qy 41 GLYGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||||||
Db 182 GGCACACGCTGGGAATGCGACAGCTGCTTTGGGATTACTGGCGGTGGCGAGACA 241
Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||||
Db 242 CTGTCTGACAGAGTGCAGAGAGCTGCTACGCTCCCAAGTCAACCAAGAACTGAGGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
    |||||||
Db 302 CTGATGACAGACCAATGAAGAGTTGAAGCCTTACAAATCGGAATGGAGGACAACATG 361
Qy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGlnGlnAlaAlaGlnAla 120
    |||||||
Db 362 ACCCGCGTACGAGAGAGAGAGCGGGCAGCGGTCTCCAAAGAGCTGCAGACGGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyLysVal 140
    |||||||
Db 422 CGGTGGCGCGGACATGAGAGAGCTGTGCGCGCCCTGTCAGTACCGCGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 482 CAGGCGCATCTCGGCCAGAGACCGAGAGCTCGGGTGGCGCTCCCTCCACCTGCGCG 541
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||||
Db 542 AAGCTCGTAAAGGCGCTCTCCCGGATCCCGATCGACCTGCAGAGAGCGCTGGAGATGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGluGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200

```

```

Db 602 CAGCGCGGCGCCCGCAGAGGCGCGCGGCTCAGCGCCATCCGACGCGCTGGGG 661
Qy 201 ProLeuValGluGlnGlnLysArgValArgAlaIleThrValGlySerLeuAlaGlnInpo 220
    |||||||
Db 662 CCCCTGTGGAAACAGGGCGCGGTGGCGGCGCCACTGTGGCTCTCCCTGGCCGCGCAGCG 721
Qy 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnGlnGly 240
    |||||||
Db 722 CTACAGAGCGGGCCCGAGGCTGGGGGAGCGGCTGCGCGCGGATGAGAGGATGGCG 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||||
Db 782 AGTCGAGCCCGCAGCCGCTGACGAGTGAAGACAGTGCAGTGGCGGCGCGCGCAAG 841
Qy 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
    |||||||
Db 842 CTGAGAGACAGGCGCCAGCAGATACCTGCGAGGCGGAGCCCTCCAGGCCGCTCAG 901
Qy 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
    |||||||
Db 902 AGCTGTTGAGGCGCTGGTGGAGACATGACAGCCGACGTGGCGCGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||||
Db 962 GTGCAGGCTGCGCTGGGACACACGCGCCGCTGTGCCAGGACACATAC 1012

```

RESULT 2

US-08-949-155-5

Sequence 5, Application US/08949155

Patent No. 6271436

GENERAL INFORMATION:

APPLICANT: Piedrahlta, Jorge A

TITLE OF INVENTION: Compositions and Methods for the

NUMBER OF SEQUENCES: 51

ADDRESS: ARNOLD, WHITE AND DURKEE

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: US

ZIP: 77210-4433

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/949,155

FILING DATE: Concurrently Herewith

CLASSIFICATION: 800

Prior APPLICATION DATA:

APPLICATION NUMBER: US 60/027,338

FILING DATE: 11-OCT-1996

Prior APPLICATION DATA:

APPLICATION NUMBER: US 60/046,094

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Hibler, David W.

REGISTRATION NUMBER: 41,071

REFERENCE/DOCKET NUMBER: TAAK:177

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

TELEFAX: (713) 789-2679

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1126 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:


```

QY 38 TTPGlnSerGlyGlnArGTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGAGGAGCCAGCCCTGGGAGAGCCCTGGCCGCTTCTGGATTACCTGCCCTGG 218
QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGCAGTCCCTGTCTCAAGTCAGAGGAGAGACTCTCCAGCAAGGACACCCAGGAA 278
QY 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATAGAGAGAGACATGAAGAGAGTGAAGCCCTACCGCGAGAGCTGGAG 338
QY 98 GluGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCGCAGCTGGCCCTGTGACCCAGGAGAGCGCCGCTGTCCAAAGAGCTGAGCGGC 398
QY 118 AlaGlnAlaArgLeuGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCGCAGGAGCCGCTGGGCGCCGACATGAGAGAGCTGCGCAACCGCTTGGTCTTACCGC 458
QY 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCGAGGTGACAAACATGTTGGGCGACACCGAGAGAGCTGCGGAGCGCGCTTCC 518
QY 158 HisLeuArgLysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTGCGGAGAGCTGCGCAAGCGCGCTGCTCCGCGACACCGAGAGACTTGGAGAGCCCTG 578
QY 178 AlaValTyrGlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGlu 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCGGTATACAGAGCGGCGCTGCGCGAGGCGCGCGACGCTGAGCGCCCTCCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CCGCTGGGCGCCCTGTGTGAGAGAGCGCATTCGCGCGCGACCTGAGTACCAAGGCGC 698
QY 218 GlyGlnProLeuGlnGluArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGlu 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GCGCAGCCCTGCGGAGGCGCGGGAAGCTTGGGCGCAAGCTTGGCGCGAGCTGGAG 758
QY 238 GluMetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluVal 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GAGATGGCAGCGCGAGCCCGCGACCGCTGTGATGATGCTGTGAGAGAGGAGG 818
QY 258 ArgAlaLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CCGACCAAGTGGAGAGAGGAGGAGGACAGTTGCGCTGAGGCGCGAGGAGATTCCACGCC 878
QY 278 ArgLeuLysSerThrPheGluProLeuValGluAspMetGlnArgGlnThrAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCCTCAAGGCGTGTTCAGAGCTCTGTGTGAGAGACATACGGGCGCGAGGCGCGCTG 938
QY 298 ValGlnThrValGlnAlaAlaValGly---ThrSerAlaAlaProValProSerAspAsn 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGGAGAGAGTGCAGTGGCGCGGTGAGACATAGCTCTCCACCTCTGCGCCAGTGAAT 998

```

RESULT 4
US-08-949-155-51

; Sequence 51, Application US/08949155

; Patent No. 6271436

; GENERAL INFORMATION:

; APPLICANT: Piedrahita, Jorge A

; APPLICANT: Bazer, Fuller W

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Generation of Transgenic Animal Species

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESS: ARNOLD, WHITE AND DURKEE

; STREET: P. O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: US

; ZIP: 77210-4433

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hilder, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-949-155-51

```

Alignment Scores:
Pred. No.: 4 87e-89 Length: 4267
Score: 998.00 Matches: 217
Percent Similarity: 58.43% Conservative: 43
Best Local Similarity: 48.76% Mismatches: 46
Query Match: 62.97% Indels: 140
DB: 4 Gaps: 5

```

US-09-827-854-19 (1-317) x US-08-949-155-51 (1-4267)
QY 1 MetLysValLeuThrPalaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGGCTGTGTAAGTGGCTCGT-----CGACGATGCGGAGACAGAGAGC 2489
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCGCGAG-----CCGCGCGGAGAGTGCACGTGTGAGGAGAGAGCCCAAG 2537
QY 38 TTPGlnSerGlyGlnArGTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGGAGCCAGCCCTGGGAGAGGAGCGCCCTGTGTGAGATTAACCTGGCTGG 2597
QY 58 ValGlnThrLeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGCAGTCCCTGTCTGACCAAGTGCAGAGAGAGACTGTCTCAGCACCAAGTCCACCAAGAA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGGCCACCGACCTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGCGAAGCGTGTCTGTGAGCCCTGAGCTCCACCGCTTCGGGTTCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTTGTGCGCACTTGTGGGGGTCTGGGTCTGTCTTTTCTTTTCTTCTTCTTCTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGGAAGAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTCC 2896
QY 78 ----- 78

```


QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 482 CAGGGCATGCTGGCGCAGAGCAACGAGAGCTGGGTCGCCCTCCCTCCACCTGCCGCC 541
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 542 AAGCTGCTGTAAGCGGCTCTCCGCGATCCCGATGACTGCAAGAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGluValArgGlyLeuSerAlaIleArgGluLysArgLeu 199
|||||
Db 602 CAGGCGCGGGGCCGCGAGCGCCGAGCGGCTCTCAGCGCATCCGCGAGCGGCTGG 658
RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5525493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road,
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22133-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/849, 389
; FILING DATE: 19920519
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16787/168/DFBC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: HUMAN LIPOPROTEIN E GENE
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..330
; US-07-849-389-6
Alignment Scores:
Pred. No.: 4.96e-42 Length: 330
Score: 507.00 Matches: 106
Percent Similarity: 98.18% Conservative: 2
Best Local Similarity: 96.36% Mismatches: 2
Query Match: 31.99% Indels: 0
Gaps: 0
US-09-827-854-19 (1-317) x US-07-849-389-6 (1-330)
QY 87 LysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeuThrProValAlaGluGlu 106
|||||
Db 1 AAGGAGTTGAAGGCTTACAAATCGAATCGAGGAGAACACTGACCCCGTGGCGGAGAG 60

QY 107 ThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMet 126
|||||
Db 61 ACGCGGCGACGGGCTGTCCAGAGAGCTGCAGCGGCGGAGGCCCGCTGGGCGCGACATG 120
QY 127 GluAspValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGln 146
|||||
Db 121 GAGGAGCTGCGGCGCGCTGTGTCAGTACCGCGGAGAGGTGACAGGCCATGCTCGGCCAG 180
QY 147 SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeu 166
|||||
Db 181 AACACCGAGAGACTCGGGGTGGCGCTCGCCCTCCACCTCGGCAAGCTGCTAAGCGCTC 240
QY 167 LeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlu 186
|||||
Db 241 CTCGCCGATGCCGATGACTGTCAGAAAGCGCTGGCAGTGTACAGGCCCGGCGCGAG 300
QY 187 GlyAlaGluArgGlyLeuSerAlaIleArg 196
|||||
Db 301 GCGCGCGAGCGCGGCTCTCAGCGCATCCGCC 330
RESULT 8
US-08-617-256-24
; Sequence 24, Application US/08617256
; Patent No. 6043031
; GENERAL INFORMATION:
; APPLICANT: Ryster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617, 256
; FILING DATE: March 18, 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: March 17, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth A.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: SQI-013CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-617-256-24
Alignment Scores:
Pred. No.: 1.2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
Gaps: 0
US-09-827-854-19 (1-317) x US-08-617-256-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
Db 2 GCACGGCTGTCGAAGAGCTGCAGCGCGCCAGCGCCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCCATGCTCGCGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
Db 121 GAGGAGCTCGGGTGGCTCGCTCGCTCCACCTGCGCAAGCTGGTAAAGCGCTCTCCCG 180
QY 169 AspaAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
Db 181 GATGCCGATGACTGCAGAAAGTCCCTGGCAGTGTACACAGCGCGGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
|||||
Db 241 GAGCGGGGCTC 252
RESULT 9
US-09-287-141-24
; Sequence 24, Application US/09287141
; Patent No. 6197498
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,141
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-287-141-24
Alignment Scores:
Pred. No.: 1.2e-28 Length: 252
Score: 370.00 Matches: 81

Percent Similarity: 97.628 Conservative: 1
Best Local Similarity: 96.438 Mismatches: 2
Query Match: 23.34% Indels: 1
DB: 4 Gaps: 0
US-09-827-854-19 (1-317) x US-09-287-141-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCGAAGAGCTGCAGCGCGCCAGCGCCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCCATGCTCGCGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg 168
|||||
Db 121 GAGGAGCTCGGGTGGCTCGCTCGCTCCACCTGCGCAAGCTGGTAAAGCGCTCTCCCG 180
QY 169 AspaAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
Db 181 GATGCCGATGACTGCAGAAAGTCCCTGGCAGTGTACACAGCGCGGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
|||||
Db 241 GAGCGGGGCTC 252
RESULT 10
US-09-431-613-24
; Sequence 24, Application US/09431613
; Patent No. 6221601
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:

Pred. No.:	1.2e-28	Length:	252
Score:	370.00	Matches:	81
Percent Similarity:	97.62%	Conservative:	1
Best Local Similarity:	96.43%	Mismatches:	2
Query Match:	23.34%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-19 (1-317) x US-09-431-613-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGAGCCCGGCTGGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGCGCGCGAGGTCATGCTCGCGCAGAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
|||||
DB 121 GAGGAGCTGGGGGTGGCTGCTCCACCTGCCAGAGCTGCTGAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGAGCGGGGCCCGGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605

GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:

Pred. No.:	1.2e-28	Length:	252
Score:	370.00	Matches:	81
Percent Similarity:	97.62%	Conservative:	1
Best Local Similarity:	96.43%	Mismatches:	2
Query Match:	23.34%	Indels:	1
DB:	4	Gaps:	0

US-09-827-854-19 (1-317) x US-09-504-245-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAGAGAGCTGCAGCGCGCGAGCCCGGCTGGCGGACATGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGCGCGAGGTCAGGCTGCGCGCAGAGACAC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
|||||
DB 121 GAGGAGCTGGGGGTGGCTGCTCCACCTGCCAGAGCTGCTGAGCGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACAGAGCGGGGCCCGGAGGCGCC 240

QY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478

GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

```

? TELEPHONE: 619-450-8400
? TELEFAX: 619-587-5360
? TELECOMMUNICATION INFORMATION
? TELEPHONE: (617)227-7400
? TELEFAX: (617)227-5941
? INFORMATION FOR SEQ. ID NO: 24:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 252 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
?
US-09-287-682-24

```

Alignment Scores:	
Pred. No.:	1,2e-28
Length:	252
Score:	370.00
Percent Similarity:	97.62%
Best local Similarity:	96.43%
Query Match:	23,348
DB:	4
Gaps:	0
Mismatches:	2
Indels:	1
Matches:	81
Conservative:	1
Length:	252

US-09-827-854-19 (1-317) x US-09-287-682-24 (1-252)

QY	109	AlAaTrrgeuSeRrYsgIueGInAlaIaGInAlaTrrgeuGlyAlaAspMeTgIuaSp	128
Db	2	GCACGGCTGTCCAAGAGACTGCAGGCGGCGGAGCCCGCTGGCGCGGCAATGAGAGAC	61
QY	129	ValCysGclAtrGleuValGlnTrrAtrGlyIuValGlnAlaMetIueGlyGlnSerTrh	148
Db	62	GTCGTGGCC - CCGCTGTGTGCAGTACCGCGGAGAGGTGCAGGCCCATGTCTCGGCCACAGACAC	120
QY	149	GInGInIueAtrGValAtrGleuAlaSerHisIueAtrGlyIueAtrGInAtrGleuAlaTrg	168
Db	121	GAGGAGGCTGGCGGGTGGCCCTTGCTCCACCTGGCAAGTGGTAAAGCGGCTCTCCG	180
QY	169	AspAlaAspAspIueGInIlyAtrGleuAlaValAtrGInAlaGlyAlaAtrGInGlyAla	188
Db	181	GATCGCCGATGACGTGCAGAAAGTCCCTGGCAGTGAACAGCGCGGGGCCGAGGGCGCC	240
QY	189	GluAtrGclGleu	192
Db	241	GAGCGCGGCTTC	252

RESULT 13
US-09-287-679-24

```

: GENERAL INFORMATION:
:
: APPLICANT: K ster, Hubert
:
: APPLICANT: little, Daniel P.
:
: APPLICANT: Braun, Andreas
:
: TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
:
: NUMBER OF CLAIMANTS: 21

```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: 08/406,199
3      FILING DATE: 03-MAR-1995
4      ATTORNEY/AGENT INFORMATION:
5      NAME: Seidman, Stephanie L
6      REGISTRATION NUMBER: 33,779
7      REFERENCE/DOCKET NUMBER: 24736-2002C
8      TELECOMMUNICATION INFORMATION:
9      TELEPHONE: 619-450-8400
10     TELEFAX: 619-587-5360
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE: (617)227-7400
13     TELEFAX: (617)227-5941
14     INFORMATION FOR SEQ ID NO: 24:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH: 252 base pairs
17     TYPE: nucleic acid
18     STRANDEDNESS: single
19     TOPOLOGY: linear
20     MOLECULE TYPE: CDNA
21     IS-Org-287-679-24

```

Alignment scores:	
Pred. No.:	1,2e-28
Score:	370.00
Percent Similarity:	97.62%
Best local Similarity:	96.43%
Query Match:	23,334
DB:	4
Length:	255
Matches:	81
Conservative:	2
Mismatches:	1
Indels:	1
Gaps:	0

US-09-827-854-19 (1-317) x US-09-287-679-24 (1-252)

QY	109	ALLATrgLeuSerLysGSLuLeuGlnAlaIaAGlnAlaIaArgLeuGlyAlaAspMetGluAsp	128
Db	2	GCACGGCTGTCCAGAGAGCTGCAGGGGGCCAGGCCGCTGGGGCCGGACATGGAGAC	61
QY	129	ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr	148
Db	62	GTGTGGCC - CGCTCGTGTGCATGCCGGAGAGGTGCAGGCTCATGTCTCGGCCAGACACC	120
QY	149	GluGluLeuArgValATArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuLeuArg	168
Db	121	GAGAGCTGGCGGGTGGCTTCCTCCCACTGGCGAAAGCTGGGTAAAGCGGCTTCCTCCG	180
QY	169	AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla	188
Db	181	GATCCCGATGACGTGCAGAAAGTCCCTGGCAGTGCACAGCGCGGGGCCCGCAGAGGGCC	240
QY	189	GluArgGlyLeu	192
Db	241	GAGCGCGGGCTC	252

RESULT 14
US-09-397-766-24

GENERAL INFORMATION:
APPLICANT: K sier, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
K. Sier, Hubert K. Sier, 1000 N. Main Street, Suite 100, Chicago, Illinois 60610

APPLICATION NUMBER: US/09/397,766
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-20021
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-397-766-24

Alignment Scores:
Pred. No.: 1,2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-19 (1-317) x US-09-397-766-24 (1-252)

OY 109 ALaATgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
|||||
DB 2 GCACGGCTGTCGAAGAGCTGCAGGCGGCGGCCGCTGGCGGACATGAGAGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGGCGGCGAGTGCATGCTGCGCCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
|||||
DB 121 GAGGAGCTGCGGCTGCGCTGCCCTCCACCTGCGCAAGCTGCTAAAGCGGCTCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGCGGCGGCCGCGAGGCGCC 240
OY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

RESULT 15

US-09-287-681-24
Sequence 24, Application US/09287681
Patent No. 6277573

GENERAL INFORMATION:
APPLICANT: K Ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,681
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-681-24

Alignment Scores:
Pred. No.: 1,2e-28 Length: 252
Score: 370.00 Matches: 81
Percent Similarity: 97.62% Conservative: 1
Best Local Similarity: 96.43% Mismatches: 2
Query Match: 23.34% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-19 (1-317) x US-09-287-681-24 (1-252)

OY 109 ALaATgLeuSerLySGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
|||||
DB 2 GCACGGCTGTCGAAGAGCTGCAGGCGGCGGCCGCTGGCGGACATGAGAGAC 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
|||||
DB 62 GTGTCCGC-CGCTGTGTGAGTACCGGCGGCGAGTGCATGCTGCGCCAGAGCAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArg 168
|||||
DB 121 GAGGAGCTGCGGCTGCGCTGCCCTCCACCTGCGCAAGCTGCTAAAGCGGCTCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
|||||
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGCGGCGGCCGCGAGGCGCC 240
OY 189 GluArgGlyLeu 192
|||||
DB 241 GAGCGCGGCTC 252

Search completed: March 14, 2003, 20:18:56
Job time : 39.7343 secs

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic acid search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585
Sequence: 1 MKVLMALLVTFAGCCQAKV.....YEKVOAAVGTSAAPVSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame.p2n.model -DEV=xlp
-DB=pubmed_Application_NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT-0 -UNITS=dits -START=1 -END=1 -MATRIX=DLSUM62
-TRAN=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=ext -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854 -JCGN1.1.298 -JCGN2.1.298 -JCGN3.1.298
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MAP -LARGEDEPT -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARM_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCF_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCFUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1585	100.0	1156	10	US-09-827-854-12
2	1581	99.7	1156	9	US-09-870-759-129
3	1581	99.7	1156	9	US-09-802-640-17
4	1581	99.7	1156	10	US-09-827-854-8

5	1581	99.7	1291	12	US-10-044-090-454	Sequence 454, App
6	1573	99.2	1156	10	US-09-827-854-9	Sequence 9, Appl
7	1573	99.2	1156	10	US-09-827-854-11	Sequence 11, Appl
8	1572	99.2	1157	10	US-09-954-456-760	Sequence 760, App
9	1572	99.2	1157	10	US-09-880-107-2244	Sequence 2244, App
10	1569	99.0	1156	10	US-09-827-854-7	Sequence 7, Appl
11	1566	98.8	1156	10	US-09-827-854-10	Sequence 10, Appl
12	1365.5	86.2	41907	10	US-09-967-013-5	Sequence 5, Appl
13	867.5	54.7	786	10	US-09-925-302-133	Sequence 133, App
14	625	39.4	478	10	US-09-964-824-374	Sequence 374, App
15	625	39.4	478	10	US-09-880-107-2491	Sequence 2491, App
16	459	29.0	356	10	US-09-960-352-5420	Sequence 5420, App
17	458.5	28.9	449	10	US-09-960-352-4726	Sequence 4726, App
18	439.5	27.7	444	10	US-09-960-352-1395	Sequence 9395, App
19	435	27.4	412	10	US-09-960-352-8237	Sequence 4237, App
20	428.5	27.0	416	10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.7	423	10	US-09-960-352-14047	Sequence 14047, A
22	405.5	25.6	409	10	US-09-960-352-5148	Sequence 5148, App
23	401.5	25.3	416	10	US-09-960-352-9797	Sequence 9797, App
24	397	25.0	253	10	US-09-179-5368-130	Sequence 130, App
25	387.5	24.4	425	10	US-09-960-352-3497	Sequence 3497, App
26	380.5	24.0	390	10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391	10	US-09-960-352-1278	Sequence 1278, App
28	379.5	23.9	392	10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401	10	US-09-960-352-1187	Sequence 4187, App
30	375.5	23.7	413	10	US-09-960-352-8042	Sequence 8042, App
31	370	23.3	252	10	US-09-796-416-24	Sequence 24, Appl
32	370	23.3	252	10	US-09-879-341-24	Sequence 24, Appl
33	369.5	23.3	404	10	US-09-960-352-2325	Sequence 2325, App
34	369.5	23.3	405	10	US-09-960-352-14063	Sequence 14063, A
35	351.5	22.2	377	10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	353	10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378	10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	386	10	US-09-960-352-10361	Sequence 10361, A
39	348.5	22.0	388	10	US-09-960-352-11966	Sequence 11966, A
40	345.5	21.8	396	10	US-09-960-352-13457	Sequence 3457, App
41	344.5	21.7	377	10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.4	432	10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332	10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.3	376	10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.2	373	10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-827-854-12
Sequence 12, Application US/09827854
Patient No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vasiliis
APPLICANT: Kyriacos, Kyriakos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT FILING DATE: 2001-04-05
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-12
Alignment Scores:
Pred. No.: 1.67e-133 Length: 1156
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-19 (1-317) x US-09-827-854-12 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGCTGCGTGTGTCACATTCCTGACAGATGCCAGGCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPglInser 40
DB 121 GAGCAAGCGGTGGAGACAGACCGGAGCCGACCTGCGCAGACGACCGATGGTGGAGC 180
QY 41 GlyGlnArgTrpLysLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCGGAGACTGGACACTGGGTGGTTTGGGATTACTGGCGCTGGGTGGAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGGAACTGAGGGCG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGGACAGACCATGAAGAGATTGAAAGCCTTACAAATCGAGACTGGAGAACTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 361 ACCCGGTGGCGAGAGAGAGCGGCGACGGCTGTCCAGAGAGCTGCAGAGCGGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
DB 421 CGGCTGGGCGGACATGTGAGAGCTGTGGCGCGCTGTGCATGCCGCGGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCAGACACCGAGAGACTGCGGTGGCTCGCTCCGCTCCGCTGGCG 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGCTGACGGGCTCCCGCGATGCCGATGACCTGCAGAAAGCGGCGGAGGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGGCGGGGGCGGCGAGGCGCGCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220
DB 661 CCCCTGTGGACAGGCGCGCTGCGGCGCGCCACTGTGGGCTCCCTGGCGGCGCAGCGC 720
QY 221 LeuGlnGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 721 CTACAGAGAGGGGCGCAGGCGCTGGGCGAGCGGCTCGCGCGGATGGAGAGATGGGG 780
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValAlaArgAlaLys 260
DB 781 AGCGCGACCGCGACCGCTGAGCAGAGGTGAAGAGACAGGTGGCGAGGTGGCGCGCAAG 840
QY 261 LeuGlnGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
DB 841 CTGAGAGACAGGCGCCAGAGATACCTGTGCAGGCGCGGAGGCGCTTCCAGGCGCGCTCAG 900
QY 281 SerTrpPheGlnProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB 901 AGCTGTTTGAGAGCCCTGTGTGGAAGACATGCAGCGCAGTGGCGCGGCTGTGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGACAGGTGCGGTGGGACACAGCGCGCGCTGTGCTCCAGCAGCAATATCAC 1011
RESULT 2
US-09-870-759-129
; Sequence 129, Application US/09870759

; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(1014)
; OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:

Pred. No.: 3 8e-133 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-19 (1-317) x US-09-870-759-129 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGCTGCGTGTGTCACATTCCTGACAGATGCCAGGCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnThrPglInser 40
DB 121 GAGCAAGCGGTGGAGACAGACCGGAGCCGACCTGCGCAGACGACCGATGGTGGAGC 180
QY 41 GlyGlnArgTrpLysLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 481 CAGGCCATGCTCGGCCAGACACCGAGAGACTGCGGTGGCTCGCTCCGCTCCGCTGGCG 240
QY 181 GGGCAGCGCGGAGACTGGCAGCTGGGTGGTTTGGGATTACTGGCGCTGGGTGGAGACA 240
DB 181 GGGCAGCGCGGAGACTGGCAGCTGGGTGGTTTGGGATTACTGGCGCTGGGTGGAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGGAACTGAGGAGCGG 300
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 301 CTGATGGACAGACCATGAAGAGATTGAAAGCCTTACAAATCGGAATCGAGAACTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
DB 361 ACCCGGTGGCGAGAGACACCGGCGACCGGCTGTCCAAAGAGCTGCAGGCGGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnThrArgGlyGlnVal 140
DB 421 CGGCTGGGCGGACATGTGAGAGAGCTGTGGCGCGCTGTGAGTGCAGTCCGCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTCGGCCAGACACCGAGAGCTGCGGCTCGCTCCGCTCCAGCAGTGGCGC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGCTGACGGGCTCCCGCGATGCCGATGACCTTCAGAAAGCGCTGCGAGGTAC 600
QY 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 601 CAGGCGGGGGCGGAGAGGCGGCGGCGCTCAGCGCCATCCGCGAGGCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnGlnPro 220

Db 661 CCCCTGTGGAACAGGCGCGGTGGCGCCCACTGTGGCTCCCTGGCCGCGCAGCC 720
Qy LeuGlIngluArgAlaIngluAlaTropLysLysLysLysLysLysLysLysLysLys 240
Db 721 CTACAGGAGCGGCGCCAGCCCTGGGGCGAGCGGTGGCGCGCGGATGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGluValAlaGluValArgAlaLys 260
Db 781 AGCCGAGCCCGGACCGCCCTGGAGCAGGTGAAGAGCAGGTGGCGGAGTGCCGCCCAAG 840
Qy 261 LeuGlIngluGluAlaIngluAlaLeuArgLysLysLysLysLysLysLysLysLys 280
Db 841 CTGGAGGAGCAGCGCCAGACAGATACGCTGAGCGCGCGGCTTCCAGCGCCCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTGTGAGCCCTGTGTGGAAGACATGCAAGCCAGTGCGGCGGTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGTGGGACACAGCGCCGCTGTGGCCACGACAAATCAC 1011
RESULT 3
US-09-802-640-17
: Sequence 17, Application US/09802640
: Publication No. US20030036057A1
: GENERAL INFORMATION:
: APPLICANT: Braun, Andreas
: APPLICANT: Bonnal Aruna
: APPLICANT: Kleyn Patrick
: TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
: FILE REFERENCE: 24736-2048
: CURRENT APPLICATION NUMBER: US/09/802,640
: CURRENT FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 122
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapien
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (61)...(1014)
: OTHER INFORMATION: Nucleotide sequence encoding apolipoprotein E
: OTHER INFORMATION: (APOE)
US-09-802-640-17
Alignment Scores:
Pred. No.: 3.8e-133 Length: 1156
Score: 1581.00 Matches: 316
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.68% Mismatches: 0
Query Match: 99.75% Indels: 0
DB: 9 Gaps: 0
US-09-827-854-19 (1-317) x US-09-802-640-17 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGCTGCTGTGTCACATTCCTGCGCAGATGCCAGGCCAAGGTG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGGGGTGGAGACAGACCGGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGlyLysPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCCAGCGCTGGGAACCTGGCATCTTTGGGATTCACCTGGCGGTGGGTGGAGACA 240
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTGTAGCAGGTGCAAGAGAGAGCTGCTCAGTCCCGAGGTACCCAGGAAGTGAAGGCG 300

Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGluGlnLeu 100
Db 301 CTGATGGAGAGACCATTAAGAGATTGAAGCTTCAATGAGATGGAGCTGAGAGAACTG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 361 ACCCGGTGGGAGAGAACCGCGGACAGCTGTCTCAAGAGAGTGCAGAGCGCGGAGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCGCCCTGTGTGCAGTACCGGCGAGGTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCTATGCTTGGCCAGAGACACCGAGAGCTGGGGGTGGCTGTGCTCCACCTGGCG 540
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTCGCTAAGCGGCTCTCCGGATGCCATGACCTGCAGAAAGCGCTGGCAGTGAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGCGCCGAGGCGCGGAGCGCGGCTCACCGCCATCCCGAGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGGAACAGGCGCGGTGGGGCGCCCACTGTGGCTCTCTGGCGCGCCAGCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaTropLysLysLysLysLysLysLysLysLysLys 240
Db 721 CTACAGGAGCGGCGCCAGCCCTGGGGCGAGCGGTGGCGCGCGGATGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGCCCGGACCGCCCTGGAGAGGTGAAGAGCAGGTGGCGGAGTGGCGCCCAAG 840
Qy 261 LeuGlnGluGlnAlaIngluAlaLeuArgLysLysLysLysLysLysLysLysLys 280
Db 841 CTGGAGGAGCAGCGCCAGACAGATACGCTGCAAGCGCCGCTTCCAGCGCCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 AGCTGTGTGAGCCCTGTGTGGAAGACATGCAAGCCAGTGCGGCGGTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAGGCTGCGTGGGACACAGCGCGCCGCTGTGGCCACGACAAATCAC 1011
RESULT 4
US-09-827-854-8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Kyriacs, Vassilis
: APPLICANT: Zannis, Vassilis
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertiglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-827-854-8

181 Glna1ag1va1aarrag1ug1va1ag1u1arrag1vl1eussera1aj1e1arrag1u1arra1eng1v 20

Db 666 AAGCTGCGTAAGCGGCTCCTCGGCGATGCCGATGCAGAACGCCCTGCCAGTGTAC 722

Db 666 AAGCTGCGTAAGCGGCTCCTCGGCGATGCCAGTGCAGAAAGCCCTGCCAGTGTAC 722

```
Dh 756 CAGCCGGGGGGCCGGAGGGCCGAGCGGCTCAGCGCCATCCGCGAGCCCTGGG 785
Qy 201 ProleuValAGluGluGlyArgValArGlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 786 CCCCTGGTGAACAGAGGCGGCGGGGCGCCACTGTGGCTCCCTGGCCGCGCAGCG 845
Qy 221 LeuGluGluArgAlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 846 CTACAGAGAGCGGGCCGAGGCTGGGGCCAGCGGCTGGCGCGCGATGAGAGATGGGC 905
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 906 ACCCGAGCCCGCGACCGCTGGTGAACAGAGTGAAGAGAGTGGCGAGGTGGCGCCAA 965
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 966 CTGGAGAGAGCGCCAGAGATACGCTCGAGCGCGAGCGCTTCACAGCCCGCTCAAG 1025
Qy 281 SerTrpPheGluProLeuValGluLysPheMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 1026 ACCTGGTTCGAGCCCTGGTGAACAGATGACAGCCGCAAGTGGCGCGGCTGGAGAG 1085
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 1086 GTGCAAGGCTGGCGTGGGACACGCGCCGCTGTGCCAGACATCAC 1136

RESULT 6
US-09-827-854-9
; Sequence 9, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassiliis
; APPLICANT: Kypreos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-9

Alignment Scores:
Pred. No.: 1,98e-132 Length: 1156
Score: 1573.00 Matches: 315
Percent Similarity: 99.68% Conserves: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 99.24% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-19 (1-317) x US-09-827-854-9 (1-1156)
Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAGGTTCTGTGGCGTGGCTGTCTGTGCATTCCTGGCGAGATGCGACGCGCAAGTG 120
Qy 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnIleThrGluTrpGlnSer 40
Db 121 GAGCAAGGCTGGAGACAGAGCGGAGCCGAGCTGCCCGACAGACAGAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGluArgPheTrpAspTrpLysArgTrpValAlaThr 60
Db 181 GCCCAGCCCTGGGAACCTGACCTGGTGGCTTTTGGATTCCTCGCTGGGTGGACACA 240
Qy 61 LeuSerGluGlnValGluGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
```

```
Db 241 CTCTGAGACAGAGTGCAGAGAGAGCTGCTCACCTCCAGAGTACACCGAGAACTGAGGGG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluGlnLeu 100
Db 301 CTGATGGACAGAACCATGATGAAGAGTTGAAGGCTCAAAATCGGAATCGAGAGCAACAG 360
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGACGCGGCGACGCTGTCCAGAGAGCTCAGCGCGAGCGCGC 420
Qy 121 ArgLeuGlyAlaAspMetGluLysPValCysGlyArgLeuValAlaGlyTrpArgLysGlyVal 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTCCGCGCGCTGTGTCTGCTGCTGCTGCTGCTGCTG 480
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCGAGACACGAGAGAGTGGCGGTGCGCTGCCCTCCACCTGGCGC 540
Qy 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTGATAGCGGCTCTCCGCAATGCCATGCTGACAGAGTGCCTGGCAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGluGlyValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGCGCGGAGAGCGCGGCGCGGCTCAGAGCGCATCCGAGCGCGCTGGG 660
Qy 201 ProLeuValAGluGluGlyArgValArGlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAACAGAGGCGCGTGGGCGCGGCTGTGGCTCCCTGGCCGCGCAGCGG 720
Qy 221 LeuGluGluArgAlaGlnAlaThrPglGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGCGGAGGCTGGGGCGAGCGGCTGCGCGCGATGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGAGCCCGGACCGGCTGTGAGAGAGAGAGAGAGAGTGGCGGCGCGCAAG 840
Qy 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGGAGAGAGCGCCAGAGATACGCTGCGAGCGCGGCTTCACAGCCCGCTCAAG 900
Qy 281 SerTrpPheGluProLeuValGluLysPheMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
Db 901 ACCTGGTTCGAGCCCTGGTGAACAGATGACAGCCGCAAGTGGCGCGGCTGGTGAAGAG 960
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGCAAGGCTGGCGTGGGACACGCGCCGCTGTGCCAGACATCAC 1011

RESULT 7
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassiliis
; APPLICANT: Kypreos, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-827-854-11

Alignment Scores:

Pred. No.:	1,98e-132	Length:	1156
Score:	1573.00	Matches:	315
Percent Similarity:	99.68%	Conservative:	1
Best Local Similarity:	99.37%	Mismatches:	1
Query Match:	99.24%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-827-854-11 (1-1156)

```
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB ATGAAAGTTCTGTGGCTGTGCTGTGCTGACATTCTCTGACAGATGCCAGGCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB GGCACAGCGGTGAGACAGCGCGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB GGCACAGCGGTGAGACAGCGCGAGCTCGCTTTGGGATTACTGGCGTGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB CTGCTGACAGCAGGTGACAGGAGAGTGTCTACCTCCAGGTCCAGCACTGAGAACTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB CTGATGACAGACCATGAAAGAGTTGAAGCCTACAAATCGGAAGTGGAGCAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB ACCCGGTGGCGAGAGACGCGGCGACGCTGTCCAGAGCTGACAGGCGGCGAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLeuVal 140
DB CGCTGTGGCGGACATGTGAGGACGTGTGCGCGCTGTGTCAGTACCGCGCGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB CAGCCATGCTGCGGACAGACACCGAGAGCTGCGGGTGGCTCCCTCCACCTGCGG 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB AAGCTGTGAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaHisArgGluLeuGly 200
DB GGGCGGGGCGCGGACGAGGCGCGCGGCTTCAGCGCCATCCGCGAGCGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB CCCCTGTGTGAGACAGGCGCGCGGCGCGGCTGTGGCTCTCCCTGGCGGCGAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB CTACAGAGACGCGGCGCGAGGCTGTGGCGGCGGCTGTGGCGGCGAGTGGAGG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB ACCCGGACCGCGGACCGCTGTGACGAGGTGAAGGACAGGTGGCGGAGGCGCGGCAAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB CTGGAGAGACAGGCGCGAGCATACCTGTGACGCGCGAGGCGCTCCAGGCGCGCTCAG 900
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB AGCTGTGTGAGGCGCGCTGTGTGAAGACATGACGCGCAGTGGCGGCTGTGTGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
```

```
DB 961 GTCCAGGCTGCGGTGGGACACAGCGCGCGCTGTGCCAGCAGCATCATC 1011
RESULT 8
US-09-954-456-760
; Sequence 760, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 760
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-760
Alignment Scores:
Pred. No.: 2.43e-132 Length: 1157
Score: 1572.00 Matches: 314
Percent Similarity: 99.37% Conservative: 1
Best Local Similarity: 99.05% Mismatches: 2
Query Match: 99.18% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-19 (1-317) x US-09-954-456-760 (1-1157)
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB ATGAAAGTTCTGTGGCTGTGCTGTGCTGACATTCTCTGACAGATGCCAGGCAAGGTG 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThr 40
DB GGCACAGCGGTGAGACAGCGCGAGCTCGCCAGCAGACCGAGTGGCAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB GGCACAGCGGTGAGACAGCGCGAGCTCGCTTTGGGATTACTGGCGTGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB CTGCTGACAGCAGGTGACAGGAGAGTGTCTACCTCCAGGTCCAGCACTGAGAACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB CTGATGACAGACCATGAAAGAGTTGAAGGCGCTACAAATCGGAAGTGGAGCAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
```



```
; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: zannis, Vassilis
; APPLICANT: Kypros, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7
```

Alignment Scores:

Pred. No.:	4,51e-132	Length:	1156
Score:	1569.00	Matches:	315
Percent Similarity:	99.68%	Conservative:	1
Best Local Similarity:	99.37%	Mismatches:	1
Query Match:	98.99%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-827-854-7 (1-1156)

```
OY 1 MetlysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGGCTGCTGTGCTGTCACATTCCTGCAGAGATCCAGGCCAAGTGC 120
OY 21 GlnGlnAlaValAluThrGluProGluProGluLeuArgGlnGlnThrGluPglIns 40
DB 121 GAGCAAGCGGTGAGAGACAGCGGAGCCGAGCTGCGCAGACAGACCGAGTGGCAGAC 180
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGAACTGGCACTGGGTCCCTTTGGGATTACCTGCGCTGGGTGCAGACA 240
OY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGCTGAGCAGGAGTGCAGAGAGAGTGTCTCAGCTCCAGGTCAACCGAAGTGAAGGCG 300
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
DB 301 CTGATGACAGACAGCAGTGAAGAGTTGAAAGCTTACAAATCGGAACCTGGAGAACACTG 360
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGCTGGCGAGAGAGACGGCGGCAAGCTCTCCAGAGAGTGCAGGGCGGCGAGGCC 420
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
DB 421 CGGCTGGGCGCGGACATGAGAGAGTGGCGGCGCGCTGCTGACAGTACCGCGGAGAGTG 480
OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGGCGACAGACACCGAGAGACTCGGGTGGCTCCCTCCACTGCGCGC 540
OY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTCGTAAGCGGCTCCCGCGAGTCCGAGTACCTGAGAAAGCGCTGGCAGAGTAC 600
OY 181 GlnAlaGlyAlaArgGluGlnGluAlaArgGlyLeuSerAlaTyrGluArgGlyVal 200
DB 601 CAGGCGGCGCGCGCGAGGCGCGAGCGGCGGCTCAGCCCATCCCGCAGCGCGCTGGG 660
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
```

```
|||||
DB 661 CCCCTGTGGAAACAGGCGCCGCGGCGCCACACTGTGGCTCCCTGGCGGCGCACCGC 720
OY 221 LeuGlnGluArgAlaGlnAlaTyrGlyLysArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGCGCGGCGCCAGGCTGGGCGAGCGGCTGGCGCGCGGATGGAGATGGCG 780
OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGACCGCGCGCGCTGAGACAGAGTGAAGAGAGAGTGGCGGAGGTGGCGCGCAAG 840
OY 261 LeuGlnGluGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
DB 841 CTGAGAGAGAGCGCCAGCAGATACGCTCAGCGCGAGGCTTCAGCGCGCTCAG 900
OY 281 SerTrpPheGluProLeuValGluAspMetGluArgGlnTrpAlaGlyLeuValGluLys 300
DB 901 AGCTGTTTCAGCGCCCTGTGTGGAAGACATCAGCGCCAGTGGCGGCTGTGTGAGAG 960
OY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
DB 961 GTGCAGGCTGCGTGGGACACAGCGCGCGCTGTGCCAGCAGCATCAC 1011
```

RESULT 11

```
US-09-827-854-10
; Sequence 10, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
```

```
; APPLICANT: zannis, Vassilis
; APPLICANT: Kypros, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-10
```

Alignment Scores:

Pred. No.:	8,37e-132	Length:	1156
Score:	1566.00	Matches:	314
Percent Similarity:	99.37%	Conservative:	1
Best Local Similarity:	99.05%	Mismatches:	2
Query Match:	98.80%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-827-854-10 (1-1156)

```
OY 1 MetlysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGTTCTGTGGGCTGCTGTGCTGTCACATTCCTGCAGAGATCCAGGCCAAGTGC 120
OY 21 GlnGlnAlaValAluThrGluProGluProGluLeuArgGlnGlnThrGluPglIns 40
DB 121 GAGCAAGCGGTGAGAGACAGCGGAGCCGAGACTCGCGCAGACAGACCGAGTGGCAGAGC 180
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 181 GGGCAGCGCTGGGAACTGGCACTGGGTCCCTTTGGGATTACCTGCGCTGGGTGCAGACA 240
OY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGCTGAGCAGGAGTGCAGAGAGAGTGTCTCAGCTCCAGGTCAACCGAAGTGAAGGCG 300
```

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
|||||
Db 301 CGATGAGAGAGACCATGTAAGAGGCTTACAAATCGGAATCGAGAGACAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGluAlaGluAla 120
|||||
Db 361 ACCCGGTGGCGGAGAGACGGCGGACGGCTGTCCAAAGAGCTGACAGCGGGCGGCGC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgLysGluVal 140
421 CGGCTGGCGGCGACATGAGAGACGTGTGCGGCGCGCTGTGTGACAGTACCGGGCGAGTGG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 481 CAGGGCATCTCTGACACGACCGAGGAGCTGGGGTGGCGCTGCCCTCCACCTGCGCC 540
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 541 AAGCTGCGTAACGGGCTCTCCGCGATGCCGATGACCTGCAGAAAGTGCCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGluValArgLysGluLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 601 CAGGCGGGGGCGCGGAGGCGCGGCGGCGGCTGACGCGCTACCGGAGCGGCTGGGG 660
QY 201 ProLeuValGluGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGluInPro 220
|||||
Db 661 CCCGTGTGTGAACAGAGGCGCGCTGGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 221 LeuGlnGluArgAlaGlnAlaThrPheGlyGluArgLeuArgAlaArgMetGluLeuMetGly 240
|||||
Db 721 CTACAGAGAGCGGCGCGGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
|||||
Db 781 ACCCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 840
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||
Db 841 CTGGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 901 ACCTGCTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 961 GTCGAGGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1011

RESULT 12
US-09-967-013-5
; Sequence 5, Application US/09967013
; Patent No. US20020045840A1
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr, Vincent P.
; TITLE OF INVENTION: METHOD FOR GENETIC ANALYSIS OF APOE DNA
; FILE REFERENCE: 11926-022001
; CURRENT APPLICATION NUMBER: US/09/967,013
; PRIORITY FILING DATE: 2000-10-25
; PRIOR APPLICATION NUMBER: 60/206,613
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 41907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-013-5

Alignment Scores: 5.25e-112
Pred. No.: 1365.50
Score: 1365.50
Percent Similarity: 60.00%

Length: 41907
Matches: 299
Conservative: 1

Best Local Similarity: 59.80% Mismatches: 6
Query Match: 86.15% Indels: 194
Db: 10 Gaps: 1

US-09-827-854-19 (1-317) x US-09-967-013-5 (1-41907)
QY 12 PheLeuAlaGlyCysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
|||
Db 20316 TTCACACACAGAGATGTCACAGCAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 20375
QY 32 LeuArgGlnGlnThrGlnTrpGlnSerGlyGlnArgTrpGluLeuAlaLeuLysArgPhe 51
|||||
Db 20376 CTGGCGCAGACAGACGAGTGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 20435
QY 52 TrpAspTyrLeuArgTyrValGlnThrLeuSerGluGlnValGlnGluLeuLeuSer 71
|||||
Db 20436 TGGGATTAACCTCGCTGGGTGCGAGACACTGTGTGAGCAGAGTGCAGAGAGAGAGCT 20495
QY 72 SerGlnValThrGlnGluLeu----- 78
Db 20496 TCCAGAGTACCCAGAGAACTAGTGAAGTGTCCCATCTGCGCCCTTGAACCTCGTGG 20555
QY 78 ----- 78
Db 20556 GCGGCTATACCTCCCAAGTCCAGAGTTCATTCCTGCGCTGCTGCTGCTGCTGCTG 20615
QY 78 ----- 78
Db 20616 CCTGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 20675
QY 78 ----- 78
Db 20676 TGGAAATTCCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCACT 20735
QY 78 ----- 78
Db 20736 CGTCTGCT 20795
QY 78 ----- 78
Db 20796 CTCACCTGCTTCCCAAGCTGTCTTGAACCTCTTGAGGCTCAAGCATCTCTCCGCTCG 20855
QY 78 ----- 78
Db 20856 CCTCCCAAGTGTCTGGAAATTAGAGCATAGAGCACCCTTCCCGGCTCTTACTCTTCT 20915
QY 78 ----- 78
Db 20916 TCGTCTGCT 20975
QY 78 ----- 78
Db 20976 GCCTCTGCGCGCT 21035
QY 78 ----- 78
Db 21036 CCCCATCCAGGCTTCTTCCCGGCTCTCCCATCTGCGAGACCCCTCCCGGCTCTGCGGCG 21095
QY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGly 98
|||||
Db 21096 CAGGCGCTGATGAGACGAGACCATGAAAGTGAAGGCTTACAAATCGGAATCGGAGAG 21155
QY 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 118
|||||
Db 21156 ACAACTGACCCCGGTGGGAGAGAGACCGGCGGACGGCTGTCTCAAGGAGCTGCAGGCGCG 21215
QY 118 aglnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGly 138
|||||
Db 21216 GCAGGCGCGGCTGGGCGGAGAGATGAGAGACTGTGGCGCGCTGTGAGTACCGGCGG 21275
QY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaArgLeuAlaSerHis 158
|||||
Db 21276 CGAGGTGAGGCGCATGCTGGGCGAGAGACCGAGAGAGTGGGCGGCTGCGCTCTCCCA 21335

Oy	158	sleuayrglsleuaavga;glnatrgleuleuAarqspAlaplaspspleuglnltyArqleuAl	178
Db	21336	CCTGGCGAACGCTGYTAAAGCGGCTCTCCGCCATGCCATACCTGCAGAAAGYGCTGGC	21395
Oy	178	aValTrgsInalagIyaIaaArgglugIyAlaagIuaArgglyeUserAaiIeargIuar	198
Db	21396	AGTGATACCAGGCGCGGGGCCCGCGAAGGCGCCGAGCGCGGCTTCACGGCATCCGCGAGCG	21455
Oy	198	gleucllyProleuVlaIGluInglnlyArqVaIaArgIaaIaThrValIgIsertleuAlaGl	218
Db	21456	CTGGGGGCCCTCGGTGGGAACAGGGCGGCTGGCGGGCCGCGCACTGTGGGCTCCTTG6CCG	21515
Oy	218	yGlnProlenuglngIuaArgAlaagInalAleTpdlYgluaTrgleuAarqAlaArqMetGIUGl	238
Db	21516	CCAGCGCGCTACAGAGAGCGGGGCCCGCACGGCCTGGGGCCAGCGCGCTCGCCCGCGGATGGAGA	21575
Oy	238	umetClSerArqThraArgaspArgIleuaspluValLysglugInValAlaGlUuaIaAr	258
Db	21576	GATGGGCGAGCGGAGCCCCCGACCGCTGGAGAGSTGAAGGAGCAGGTGGCGGAGGTGG	21635
Oy	258	galAyalsleuclugIugInlalagInglnllearygleuclInalaglualaphngInAlaAr	278
Db	21636	CGCAAAGCTGGAGAGAGAGGCCCGCACAGATATACGCTGCAGGCGGAGGCGCTTCCAGGCGCG	21695
Oy	278	gleuYsserTrppheglUpProleuValGIuaspmetGlnarqgInArqgIntRpAlaGlYleuVa	298
Db	21696	CCTCAAGACTGTGTTCCGAGCCCTCTGTGGTAAGACATGCACAGCGCCAGTGGCGGCGTGGT	21755
Oy	298	IglulysValIglualaIalaValaIglYThrSeraIaIaIaProvalProserAspaSnHIs	317
Db	21756	GGAGAAAGTGACAGCTGCGCTGGCGACCAAGCGCGCCCTGTGGCCACGACAATCAC	21813
RESULT 13			
US-09-925-302-133			
:	:	: Sequence 133, Application US/09925302	:
:	:	: Patent No. US20020044941A1	:
:	:	: GENERAL INFORMATION:	:
:	:	: APPLICANT: Rosen et al.	:
:	:	: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies	:
:	:	: FILE REFERENCE: PA104	:
:	:	: CURRENT APPLICATION NUMBER: US/09/925,302	:
:	:	: CURRENT FILING DATE: 2001-08-10	:
:	:	: PRIOR APPLICATION NUMBER: PCT/US00/05918	:
:	:	: PRIOR FILING DATE: 2000-03-08	:
:	:	: PRIOR APPLICATION NUMBER: 60/124,270	:
:	:	: PRIOR FILING DATE: 1999-03-12	:
:	:	: NUMBER OF SEQ ID NOS: 896	:
:	:	: SOFTWARE: PatentIn Ver. 2.0	:
:	:	: SEQ ID NO 133	:
:	:	: LENGTH: 786	:
:	:	: TYPE: DNA	:
:	:	: ORGANISM: Homo sapiens	:
:	:	: US-09-925-302-133	:

Alignment Scores:			
Pred. NO.:	1,79e-69	Length:	786
Score:	867.50	Matches:	183
Percent Similarity:	98.39%	Conservative:	0
Best local Similarity:	98.39%	Mismatches:	3
Query Match:	54.73%	Indels:	2
DB:	10	Gaps:	0
US-09-827-854-19 (1-317) x US-09-925-302-133 (1-786)			
QY	132	ArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlnInserrHrgIugLueu	151
Db	6	CGCCTGTCAGTACCGGGGAGGTGCACACCATGCTCGGCCAGACACCGAGACTG	65
QY	152	ArgValArgLeuAlaSerHisLeuArgLysLeuArgGlnArgLeuArgAspAlaAsp	171
Db	66	CGGGTGGCCCTCGGCTCCACATCTGGCA-CTGGTAA-CGGCTCCCTCCGATGCCGAT	123

OY	172	AsLeuGlnLysArGLeuAlaValTyrGlnAlaIleAlaArgGluGluGlyValArgGly	191
Db	124	GACCTCGAAGAGCCCTTGCAAGTACCAGGCCGGGGCCCGCAGAGGCCGACGGCGC	183
OY	192	LeuSerAlaIleArGValArGLeuGlyProLeuValGluGlnGlyArGValArgAlaAla	211
Db	184	CTCACGCCCATCCCGAGAGCCCTGGGGCCCTCGTGGAGAACAGGGCCCGCGCGGGCGCC	243
OY	212	ThrValGlySerLeuAlaGlnGlnProLeuGlnGluArgAlaGlnAlaTyrGlyValArg	231
Db	244	ACTGGGGCTCTCCGGCCGGCCACCCCTACAGAGCGGGCCAGAGCCTGGGGCGAGCGG	303
OY	232	LeuArGAlaArGMeGluGluMeGlySerArGThrArGAspArgLeuAspGluValLys	251
Db	304	CTGGCGCGCGGATGGAGAGATGGGAGCGGAGCCCGCGACCGCCTGCAGCAGAGTGAAAG	363
OY	252	GluGlnValAlaGluValAlaArgAlaLysLeuGluGluGlnAlaGlnGlnIleArGLeuGln	271
Db	364	GAGCAGGGTGGGAGGAGTGGCGCCAACTGGAGAGCAGAGCCGACGATACCCCTGGCAC	423
OY	272	AlaGluAlaPheGlnAlaArgLeuLysSerTyrPheGluProLeuValGluAspMetGln	291
Db	424	GCCGAGGCTTCCAGGCGCCGCTAAAGCTGGTTCAGGCCCTCGTGGAAAGACATGCAC	483
OY	292	ArgGlnTrrAlaGlyLeuValGluLysValGlnAlaAlaValGlyThrSerAlaAlaPro	311
Db	484	CGCCAGTGGCGCGGGCTGGTGGAAGAAAGTCAGAGCTCCCTGGGGACACAGCCGCCCCCT	543
OY	312	ValProSerAspAsnHis	317
Db	544	GTGCCACAGCACAATTCAC	561

```

RESULT 14
US-09-964-824A-374/c
: Sequence 374, Application US/09964824A
: Patent No. US20020102531A1
:
: GENERAL INFORMATION:
:   APPLICANT: Horrigan, Stephen
:   TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
:   TITLE OF INVENTION: Sets
:   FILE REFERENCE: 689290-73
:   CURRENT APPLICATION NUMBER: US/09/964, 824A
:   CURRENT FILING DATE: 2001-09-27
:   PRIOR APPLICATION NUMBER: US/60/236,033
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,032
:   PRIOR FILING DATE: 2000-09-28
:   PRIOR APPLICATION NUMBER: US/60/236,028
:   PRIOR FILING DATE: 2000-09-28
:   NUMBER OF SEQ ID NOS: 583
:   SOFTWARE: PatentIn version 3.0
:   SEQ ID NO 374
:   LENGTH: 478
:   TYPE: DNA
:   ORGANISM: Homo sapiens
:   FEATURE:
:   NAME/KEY: misc_feature
:   LOCATION: (1)---(478)
:   OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

```

Alignment Scores:		
Pred. No.:	5, 066-48	478
Score:	625.00	126
Percent Similarity:	97.69%	
Best Local Similarity:	96.92%	
Query Match:	39.43%	
DB:	10	0
US-09-827-854-19 (1-317)	x US-09-964-824A-374 (1-478)	
QY	1 MetLysValLeuTyrAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20

Db 396 ATGAAGTTCTGTGGGCGCTGCTGTGCACATTCTGGCAGATGCCAGCCCAAGGTG 337
Qy 21 GtGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCAGCTGCGCAGACAGCCAGTGGCAGAC 277
Qy 41 GtGlnAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGGTGGAGACAGCTGCTGCTTTTGGATTAACCTGGGCTGGTGCAGACA 217
Qy 61 LeuSerGluGlnValaGlnGluLeuLeu-SerSerGlnValaThrGlnGluLeuArgAl 80
Db 216 CTGCTGACAGCTGAGAGAGAGCTGCTGAGCTCCAGCTCCAGCCAGAGACTGAGGCG 157
Qy 80 aleuMetAspGluThrMetLeuGlyAlaTyrTrpSerGluLeuGlnGluLeu 100
Db 156 GCTGATGACGAGACCATTAAGAGCTTGAAGCCTTACAAATGGAATGGAGAACACT 97
Qy 100 uThrProValaAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
Db 96 GACCCCGGTGGCGAGAGAGACCGGCGCAGCTGTCCAAGAGCTGCAAGGCGCGGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGACATGAGAGACGTG 9

RESULT 15

US-09-880-107-2491/c
; Sequence 2491, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2491
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
; NAME/KEY: unsure
; LOCATION: (1)..(478)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

Alignment Scores:

Pred. No.:	5,06e-48	Length:	478
Score:	625.00	Matches:	126
Percent Similarity:	97.69%	Conservative:	1
Best Local Similarity:	96.92%	Mismatches:	2
Query Match:	39.43%	Indels:	1
DB:	10	Gaps:	0

US-09-827-854-19 (1-317) x US-09-880-107-2491 (1-478)

Qy 1 MetLysValLeuTrpAlaAlaLeuValaThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 336 ATGAAGTTCTGTGGGCGCTGCTGTGCACATTCTGGCAGATGCCAGCCCAAGGTG 337
Qy 21 GtGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGAGCCGAGCCNCAGCTGCGCAGACAGCCAGTGGCAGAC 277

Qy 41 GtGlnAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACAGCGGTGGAGACAGCTGCTGCTTTTGGATTAACCTGGGCTGGTGCAGACA 217
Qy 61 LeuSerGluGlnValaGlnGluLeuLeu-SerSerGlnValaThrGlnGluLeuArgAl 80
Db 216 CTGCTGACAGCTGAGAGAGAGCTGCTGAGCTCCAGCTCCAGCTCCAGCCAGAGACTGAGGCG 157
Qy 80 aleuMetAspGluThrMetLeuGlyAlaTyrTrpSerGluLeuGlnGluLeu 100
Db 156 GCTGATGACGAGACCATTAAGAGCTTGAAGCCTTACAAATGGAATGGAGAACACT 97
Qy 100 uThrProValaAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
Db 96 GACCCCGGTGGCGAGAGAGACCGGCGCAGCTGTCCAAGAGCTGCAAGGCGCGGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGGCGGAGACATGAGAGACGTG 9

Search completed: March 15, 2003, 03:05:08
Job time : 80.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds
(without alignments)

3980.771 Million cell updates/sec

Title: US-09-827-854-19

Perfect score: 1585

Sequence: 1 MKVLMALVTFLLAGCOAKV.....VEKVOAAVGSAPVPSDNH 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/rtnat_11032003_101610_27495/app_query.fasta.1.3576
-DB=EST -QFMT=fastap -SUFMT=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn.1.1.8826 -runat_11032003_101610_27495 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST :
1: em_estba :
2: em_esthum :
3: em_estlin :
4: em_estmu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_hic :
9: gb_est1 :
10: gb_est2 :
11: gb_hic :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: gb_gss :
18: em_gss_hum :
19: em_gss_inv :
20: em_gss_pln :
21: em_gss_vit :
22: em_gss_fun :
23: em_gss_mam :
24: em_gss_mus :
25: em_gss_other :
26: em_gss_pro :
27: em_gss_rod :

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1286.5	81.2	1027	13	BI670367
2	1250	78.9	842	12	BG763371
3	1244	78.5	933	13	BI668318
4	1216	76.7	938	12	BG761746
5	1213.5	76.6	942	13	BI600906
6	1199	75.6	800	13	BM042094
7	1198.5	75.6	922	12	BG472299
8	1195.5	75.4	922	13	BI597743
9	1186.5	74.9	817	12	BG774871
10	1185	74.8	811	13	BI600563
11	1171.5	73.9	845	12	BG829472
12	1148	72.2	790	12	BG707147
13	1145	72.2	919	13	BI551475
14	1135	71.6	757	13	BM042228
15	1125.5	71.0	907	12	BG706129
16	1124	70.9	706	14	BM728696
17	1124	70.9	741	12	BG762924
18	1122	70.8	1110	11	AK010261
19	1120	70.7	804	12	BG702752
20	1119	70.6	834	13	BM042676
21	1116.5	70.4	808	13	BI668329
22	1113	70.2	855	13	BI613362
23	1110	70.0	803	13	BI670350
24	1107	69.8	782	12	BG716776
25	1107	69.8	812	13	BI601551
26	1107	69.8	965	14	BQ677266
27	1104	69.7	757	13	BI603658
28	1101.5	69.5	797	12	BG715366
29	1101	69.5	746	13	BI159757
30	1100	69.4	846	13	BM042153
31	1095	69.1	757	12	BG707750
32	1095	69.1	794	13	BI601279
33	1095	69.1	954	12	BF967543
34	1093	69.0	790	13	BI551066
35	1092	68.9	790	13	BI551811
36	1088.5	68.7	798	12	BG708414
37	1087.5	68.6	891	13	BI548292
38	1081.5	68.2	802	13	BI458355
39	1080	68.1	1100	14	BM914382
40	1079.5	68.1	914	13	BI603523
41	1079	68.1	748	13	BI553085
42	1078.5	68.0	930	12	BF967857
43	1074.5	67.8	812	12	BG769968
44	1074	67.8	688	13	BI544886
45	1073.5	67.7	803	12	BG709360

ALIGNMENTS

RESULT 1
BI670367
LOCUS 603292738F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',
DEFINITION B1670367 mRNA sequence.
ACCESSION B1670367
VERSION B1670367.1 GI:15584600
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1027)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov

Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

Plate: LLM11790 row: 1 column: 09
High quality sequence stop: 845.

FEATURES

Location/Qualifiers

1..1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 194 a 308 c 400 g 125 t
ORIGIN

Alignment Scores:

Pred. No.: 8..16e-125 Length: 1027
Score: 1286.50 Matches: 291
Percent Similarity: 90.85% Conservative: 7
Best Local Similarity: 88.72% Mismatches: 19
Query Match: 81.17% Indels: 12
DB: 13 Gaps: 2

US-09-827-854-19 (1-317) x B1670367 (1-1027)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 26 ATGAAGTTCTGTGGCTGTGTGTCACATTCCTGCGCAGAGATGCCAGCAAGTGC 85
QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnIns 40
|||||
DB 86 GACCAAGCGGTGAGACAGAGCCGAGCCGACCTCGCCAGACACCGAGTGGCAGAGC 145
QY 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnTr 60
|||||
DB 146 GCGCAGCGCTGGGAACTGCGACACTGGCTTGTGGGATTACTGGCGTGGGCGAGACA 205
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 206 CTGCTGACAGAGTGCAGAGAGTGTCTCCAGCTCCAGGTCCACGAGAACTGAGAGCGG 265
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
|||||
DB 266 CTGATGAGAGACCATGAAGAGTGAAGCCCTACAAATCGAACTGAGGAGAACTG 325
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
|||||
DB 326 ACCCGCGTGGCGAGAGACCGCGGACGCTCTCAAGAGAGTGCAGCGCGCGAGGCC 385
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 386 CGGCTGGCGCGACATGAGAGACGTGCGCGCGCTGTCAGTACCGCGCGAGGTG 445
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

DB 446 CAGGCCATGCTGGCCAGACACCGAGAGAGCTGGGCTGCCCTGCCCTCCACCTGCC 505
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaLys 180
|||||
DB 506 AAGCTCGTAAGAGGCTCTCCCGCATGCGGATGACCTCGAGAAAGCGCTGGCAGTAC 565
QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaAlaLeuGluLysLeu 200
|||||
DB 566 CAGCGCGGGGCGCCGAGGCGCGGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGG 625
QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln 220
|||||
DB 626 CCCCTGCTGGAGACAGCTCCGCTGGCGCGCCGCGACTGTGGCTCCCTGCCGCGCAGC 685
QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlu-Met 239
|||||
DB 686 GCTACAGAGAGCGGCGCAAGCGCTGGCGAGCGCTTGGCGCGCGAGTGAAGAGAG 745
QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArg 259
|||||
DB 746 GCGAGCGCGGACCGCGACCGCTGGACCGAGTGAAGAGAGAGTGGCGGAAGTGGCG 805
QY 259 LysLeu---GluGlnGlnAlaGlnGln-IleArgLeuGln-AlaGluAlaPheGlnAl 277
|||||
DB 806 CCCACCTGAGAGCGAGCCAGGCCAGCCAGACATACCTCTCAGGCGCGAGGCTTCCTAGC 865
QY 277 ArgLeuLysSerTrpPheGluPro----LeuValGluAspMetGlnArgGlnTrpAl 295
|||||
DB 866 CCGCGCTCAGAGCTGTGTGAACCCCTGGTGTGAACGACATAGGCGCGCGCAGTGGG 925
QY 295 ArgLeuValGluLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSer 315
|||||
DB 926 CGGCTGTGTGAAGAGTGCAGGCTGGCGGCGCACACCGCGCTGTGGCAGC-G 984
QY 315 spasnhs 317
DB 985 ACATCAT 992

RESULT 2
LOCUS BG763371 842 bp mRNA linear EST 15-MAY-2001
DEFINITION 602735433f1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4860585 5',
ACCESION BG763371 GI:14074024
VERSION BG763371.1
KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 842)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

TITLE

NIH-MGC http://mgi.nci.nih.gov/.
Tissue Procurement: ATCC/DC/DT/PTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory

JOURNAL

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

FEATURES

Plate: LLM1718 row: b column: 10
High quality sequence stop: 817.
Location/Qualifiers

source

1..842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4860585"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"

QY	201	ProleuvaiGlunglNglyArValAryglAlaharVaIglsyer-LeuaIaglylnPr	220
Db	657	CCCCGTGGGAACAAGGCGCCGCTGGGGGCCCGCCCACTGTGGGCTCCTTGGCCGGACACC	716
QY	220	OleuGlInuarAglaAglaIatATP-GlyGuAry-LeuaArglaAaryMetclu---Glu	238
Db	717	GCTACAGAGAGGGGGCCAGCGCTGGGGGCGAGCGGCTTGCGGGCGCGATGGAGAGCAAT	776
QY	239	MetclSyserAtythr-Arg-AspaIglyeAspGUvalllys-GluGlnVALALA--Glu	256
Db	777	TGGGACAGCCCGACCCCGCGCACCGCTTGGACAGAGCTCAAGGAGACAGTGGCGGGAAAG	836
QY	257	ValARglaALysLeuGluGlu-GlnAlaInglnlleArpreuGlnala--GlnAlaphe	275
Db	837	TTGGCGCCCAAGCTGGAAGAAGCAGCGCCAGAGATTACGCTTCAGCGCGGAGGCCTTC	896
QY	276	GlnAlaArGleuLysSerTrpPheGluProleuVal	287
Db	897	CCAGGCGCGCTCAAAGCTGCTGCAAGCCTGCTT	932
RESULT 5			
B16D0906			
DEFINITION	B16D0906	942 bp	mRNA linear EST 07-SEP-2001
LOCUS	60324924F1 NIH_MGC_96	Homo sapiens CDNA clone IMAGE:5301010 5'	
ACCESSION	B16D0906		
VERSION	B16D0906.1	GI:15493845	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	1 (bases 1 to 942)		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rgsab@remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshiyuki and Piero Carninci (RIKEN) DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LHAM1762 row: a column: 11 High quality sequence stop: 762. Location/Qualifiers 1..942 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5301010" /clone_id="NIH_MGC_96" /tissue_type="hypothalamus" /lab_host="DH10B" /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROP 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	177 a	284 c	361 g 120 t
ORIGIN			
Alignment Scores:	3 39e-117	Length:	942
Pred. No.:	1213.50	Matches:	265
Score:	93.47%	Conservative:	7
Percent Similarity:			

Best Local Similarity:	91.07%	Mismatches:	13
Query Match:	76.56%	Indels:	8
DB:	13	Gaps:	1
US-09-827-854-19 (1-317) x B1600906 (1-942)			
Oy	1 MetLysValLeuTrpAlaAlaLeuLeuVal1ThrPheLeuAlaG1CysGlnAlaLysVal	20	
Db	73 ATGAAGGTTCTCTGGGCTGCCGTTGCTGGTGCACATTCTCGGACAGAGATGCCAGGCCAAGG	132	
Oy	21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40	
Db	133 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCACAGACACCAAGTGGCGAGAGC	192	
Oy	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnTrp	60	
Db	193 GGCACGCGCTGGGAACCTGGCAGCTGGGTCTTTGGGATTACCTGGCTGGGGTGCAGACA	252	
Oy	61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnVala1ThrGlnGluLeuArgAla	80	
Db	253 CTGCTCGAGACAGTGCAGAGAGAGGCTGCTCACCTCCAGCTCACCAAGAACTGAGGGCG	312	
Oy	81 LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu	100	
Db	313 CTGAATGACAGAACCATGAAGAGTTGAAGGCTCTCAAAATCGGAATGGAGGAACAACTG	372	
Oy	101 ThrProVala1AlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120	
Db	373 ACCCGGTGGCGGAGAGAGAGCGGGGACCGGCTGTCCAAAGACTGCAGGGCGCGCAGGGC	432	
Oy	121 ArgLeuGlyAlaAspMetGluAspValaLysGlyArgLeuValaGlnTyrArgGlyGluVal	140	
Db	433 CGGCTGGGGCGCGACATGAGAGACCTGTGGCGCGCCCTGTGCAGATACCGCGCGAGGTG	492	
Oy	141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValaArgLeuAlaSerHisLeuArg	160	
Db	493 CAGGCATGAGCTGGCGACAGACACCGAGGAGCTGGGGTGGCTGCCTGCCACCTGGCGC	552	
Oy	161 LysLeuArgArgArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180	
Db	553 AAGCTGCGTAACGGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGGAGAGTAC	612	
Oy	181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAla1IleArgGluArgLeuGly	200	
Db	613 CAGCGCGGGCGCGGAGAGCGCCGACGCGCGGCTTCAGGCCATCCGACAGCGCCGCGGG	672	
Oy	201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro	220	
Db	673 CCCCTGGTGGAAACAGAGCGCGCTGGCGGGC -GCCACTGTGGCTCCCTGGCGCGGACGCC	731	
Oy	221 LeuGlnGluArgAlaGlnAlaThrPheGlyGluArgLeuArgAla - ArgMetGlu - GluMetG	240	
Db	732 CTACAGGAGCGCGGCCAGGCTGTGGGGCGAGCGCGGTGCGCGCGGATGGAGAGCAGATTG	791	
Oy	240 LysSerArgTrpThrArgAspArgLeuAsp - GluValLys - GlnGlnVala1AlaGluVal - Arg	259	
Db	792 GCAGCGGCGCGCCGACACGCGCTGGACCGAGGGTGAACGACAGGTGGCGAAGTTGGGGG	851	
Oy	259 IalLysLeuGlnGluGlnAla - - - GlnGln1IleArgLeuGlnAlaGluAlaPheGlnAla	278	
Db	852 CCAACTTGAGAGAACACAGCGCCAGCAGAAATACGCTTGACAGGCCGAGGCCCTTCAAGGCC	911	
Oy	278 rGluLeuLysSerTrpPheGluPro 285		
Db	912 GCCTCAAAAC - TGGTTGACCCCT 933		
<p>RESULT 6 BM042094 800 bp mRNA linear EST 07-NOV-2001 LOCUS BM042094 DEFINITION 60361571361 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5', mRNA sequence. BM042094 ACCESSION BM042094.1 GI:16771361 VERSION</p>			

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: DCTP/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1875 row: 1 column: 03
High quality sequence stop: 792.

FEATURES
source
1..800
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_1ib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOMB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 146 a 239 c 315 g 100 t
ORIGIN

Alignment Scores:

Pred. No.: 8 93e-116 Length: 800
Score: 1199.00 Matches: 243
Percent Similarity: 99.20% Conservative: 5
Best Local Similarity: 97.20% Mismatches: 2
Query Match: 75,658 Indels: 1
DB: 13 Gaps: 0

US-09-827-854-19 (1-317) x BM042094 (1-800)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 50 ATGAAGATTCTCTGGGCTGGCTGGTCACATTCCTGGCAGATGCCAGGCAAGGTG 109
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnPglInser 40
|||||
DB 110 GAGCAAGCGGTGGAGACAGCGCGAGCCGAGCTCGCGCAGACAGCGAGTGGCAGAGC 169
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 170 GGCACAGCTGGGAACCTGCACCTGGCTTTGGGATTACTGGCGTGGGTGCAGACA 229
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||
DB 230 CTGTCTGACAGAGTGCAGAGGAGCTGCACAGCTCCAGGTACCCAGGAACTGAGGGCG 289
QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
|||||
DB 290 CTGATGACAGAGCCATGAGAGGTGAAGGCTTACAAATCGGAACTGGAGGAAACTG 349
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
|||||
DB 350 ACCCGGCTGGCGAGAGACGCGGCGCTGTCCAGAGAGCTGCAAGGCGGCGCAGGCC 409

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 410 CGCGTGGCGCGGACATGAGGAGACGTGTGGCGCGCTGTGTGACATGCCGCGAGAGT 469
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 470 CAGGCCATGCTCGGCCAGAGACCCAGAGACTGCGGGTGGCTTCCCTCCACTTGGC 529
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 530 AACCTGGTGAAGGGCTCCCGCGAGTCCGAGTACCTGCAGAAAGGCGCTGGCAGGTAC 589
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlyArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
DB 590 CAGGCGGGCGCGCGAGGCGCGGAGCGCGCTCAGCCCATCCCGCGAGCGCTGGGG 649
QY 201 ProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
|||||
DB 650 CCGCTGGTGGACAGGGCGCGGTGGCGCCAGCACTGTGGCTCCCTGGCGGCCAGCCG 709
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGlnMetGly 240
|||||
DB 710 CTACAGAGAGCGGCGCCAGGCTGTGGGAGAGCGGTG-CGCGCGGAGTGCAGAGATGGCG 768
QY 241 SerArgThrArgAspArgLeuAspGluVal 250
|||||
DB 769 ACCGCGGAGCGCGAGCGCTGTGAGCAGCTG 798

RESULT 7
BG472299

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

EST.

ORGANISM

human.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

US-09-827-854-19 (1-317) x BM042094 (1-800)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 927)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: L1CM1419 row: k column: 08
High quality sequence stop: 848.

FEATURES
source
1..927
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_1ib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOMB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 170 a 271 c 371 g 115 t

Alignment Scores:

Pred. No.:	1.25e-115	Length:	927
Score:	1198.50	Matches:	274
Percent Similarity:	92.33%	Conservative:	3
Best Local Similarity:	91.33%	Mismatches:	13
Query Match:	75.62%	Indels:	12
DB:	12	Gaps:	1

US-09-827-854-19 (1-317) x BG472299 (1-927)

```

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB ATGAAGGTTCTGTGGCTCGTTGCTGTCACATTCTTGGCAGAGTCCAGGCAAGGTTG 106
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnPrgInser 40
DB GAGCAAGCGGTGAGACACAGCCGAGCCGAGCTGCCAGACAGCCAGAGTGGCAGAC 166
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB GGCCACCGCTGGGAACCTGCGACTGGCTTTGGATTACCTGCGCTGGGTGGCAGACA 226
QY 61 LeuSerGluGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB TTTTGTGACAGGTGACAGAGAGAGAGGTCTCACCCTCCAGGTACCCAGGAACTGAGGGCG 286
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB CTGATGAGACAGACCATGTAAGAGTTGAAGGCTTACAAATCCGAACTGAGAGACACTG 346
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB ACCCCGCTGGCGAGAGAGACGGCGGACGCTGTCCAAAGAGCTGCAGCGCGGCGAGGCC 406
QY 121 ArgLeuGlyAlaAspMetLysPvalLysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB CGGCTGGCGCGACATGAGAGCGTGCAGCGCTGTGTCACAGTACCGCGCGAGGTG 466
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnArgValArgLeuAlaSerHisLeuArg 160
DB CAGGCGATCTCGGCCAGAGACCCGAGGAGCTGGGCTGCCCTCCCACTCCGCGCC 526
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB AAGCTGCGTAAGCGGCTCTCCGCGATGCCATGACCTCAGAAAGCGCTGCGAGTGTAC 586
QY 181 Gln-AlaGlyAlaArgGlnGlyAlaGlnArgGly--LeuSerAlaAlaLeuArgGlnLeu 199
DB CAGGCGCGGGGCGCGGAGGCGCGGCGGCGCTCCTCAAGCGCAATCCGCGAGCGCCCTT 646
QY 200 --GlyProLeuValGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyG 219
DB GGGGCGCCCTGGGTGGAACAGGCGCGGCTGCCGCTGTGGCTCTGGCGCGCC 706
QY 219 InProLeuGlnLysArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlnGln 239
DB AACG-CTACAGAGCGGGCGCGGCTGGGCGGAACGCTG-CGCGGCGGAGTGAAGAGAA 764
QY 239 etGlySerArg-ThrArgAspArgLeuAspGluValLysGlnGlnValaGlnValaArg 258
DB TGGGCGAGCGGGAGCCCGGACCTGAGCGGTGAAGAGACAGGTGGCGAGGCTGG 824
QY 258 GlnAlaLysLeuGlnGlnGlnAlaGlnGlnLeuArgGlnAlaGlnAla---PheGlnAl 277
DB CCGCAAGCTGGAGAGACAGCGCCAGCAGATACGGCTGCAGGCGCGAGGGCTTCCAGGGC 884
QY 277 aArg-LeuLysSerTrpPhe-GluProLeuValGlnAspMet 290
DB CCGGCTCAAGAGCTGTGCGAGCCCTGGGTGAAGACATG 926
RESULN 8
BI597743 922 bp mRNA linear EST 07-SEP-2001
LOCUS BI597743

```

```

DEFINITION 603248609F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
ACCESSION mRNA sequence.
VERSION BI597743
KEYWORDS BI597743.1 GI:15490682
SOURCE EST.
ORGANISM human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 922)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            JOURNAL
            COMMENT
            Contact: Robert Strausberg, Ph.D.
            Email: cga@psh-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraaki
            Toshiluyki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM11760 row: b column: 04
            High quality sequence stop: 782.
FEATURES
            source
            1..922
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5300259"
            /clone.lib="NIH_MGC_96"
            /tissue.type="hypothalamus"
            /lab.host="DH10B"
            /note="Organ: brain; Vector: pBluescript (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
            size-selected for average insert size 2.3 kb and
            normalized to R07 5. This is a primary library enriched
            for full-length clones and constructed using the
            cap-trapper method (carninci, in preparation). Library
            constructed by M. Brownstein (NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 176 a 273 c 359 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 2.56e-115 Length: 922
Score: 1195.50 Matches: 256
Percent Similarity: 93.55% Conservative: 5
Best Local Similarity: 91.76% Mismatches: 14
Query Match: 75.43% Indels: 5
DB: 13 Gaps: 1
US-09-827-854-19 (1-317) x BI597743 (1-922)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB ATGAAGGTTCTGTGGCTCGTTGCTGTCACATTCTTGGCAGAGTCCAGGCAAGGTTG 132
QY 21 GluGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnPrgInser 40
DB GAGCAAGCGGTGAGACACAGCCGAGCCGAGCTGCCAGACAGCCAGTGGCAGAC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB GGCCACCGCTGGGAACCTGCGACTGGCTTTGGATTACCTGCGCTGGGTGGCAGACA 252
QY 61 LeuSerGluGlnValGlnGlnLeuLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
DB TTTTGTGACAGGTGACAGAGAGAGGTCTCACCCTCCAGGTACCCAGGAACTGAGGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
DB TTTTGTGACAGGTGACAGAGAGAGGTCTCACCCTCCAGGTACCCAGGAACTGAGGGCG 312

```


RESULT 10
BI600563
LOCUS 603244936F1 NIH_MGC_96 811 bp mRNA linear EST 07-SEP-2001
DEFINITION Homo sapiens CDNA clone IMAGE:5287329 5',
mRNA sequence.
ACCESSION BI600563
VERSION BI600563.1 GI:15493502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11726 row: 9 column: 10
High quality sequence stop: 783.
Location/Qualifiers
1..811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287329"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 148 a 245 c 316 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 2,7e-114 Length: 811
Score: 1185.00 Matches: 240
Percent Similarity: 98.37% Conservative: 2
Best Local Similarity: 97.56% Mismatches: 3
Query Match: 74.76% Indels: 1
DB: 13 Gaps: 0
US-09-827-854-19 (1-317) x BI600563 (1-811)
QY 1 MetlysaValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 74 ATGAAGGTTCTGTGGCGCTGTCTGTACATTCCTGGCAGATGCCAGCCAAAGGTG 133
QY 21 GUGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 134 GAGCAAGCGGTGGAGACAGAGCCGAGCGCCAGCCAGACCGAAGTGGCAGAGCG 193
QY 41 GUGlnArGTrpGluLeuAlaLeuGlyArGpHeTrpAspTrpLeuArGTrpValGlnThr 60
Db 194 GGCAGGCGCTGGAACTGCACTGGGTGCTTTGGGATTAACGCGCTGGGTGCAGACGA 253
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 254 CTGTCTGAGCAGGTGACGAGGAGGAGCTGTCTCCAGGTCCACCAAGGAAGTGAAGGCG 313

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 314 CTGATGAGCAGAGACCATATGAAGGCTTCAATATGGAATCGAGACCAACTGTG 373
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 120
Db 374 ACCCGGTGGGGAGAGACCGGGCAGCGCTGTCCAGAGAGCTCAGCGCGCAGGCC 433
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
Db 434 CGGTGGCGCGGACATGAGAGCTGTGCGCGCTGTGTCACTACCGCGCGAGGTG 493
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValAlaGlyAlaSerHisLeuArg 160
Db 494 CAGCGCATGCTCGGCAAGCAGCAGCAGAGAGCTCGCGGTGCGCTGCCCTCCACCTGGCG 553
QY 161 LysLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 554 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGCAGCAAGCGCCTGGCACTGTAC 613
QY 181 GlnAlaGlyAlaArgGluGluValAlaGluArgGlyLeuSerAlaIleArgGluArgLeu 200
Db 614 CAGCGCGGGCGCGGAGGCGCGGAGCGCGCTCAGCGCATCCGAGCGCGCTGGGG 673
QY 200 YProLeuValGluGlnArgValArgValAlaAlaThrValGlySerLeuAlaGlyLys 220
Db 674 CCCCTGTGTGAGACAGGCGCGCGCTGCGGCGCGCCACTGTGGGTCTCCGCGAGCAGCC 733
QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetG1 240
Db 734 GCTACAGAGAGCGGCGCGCGCTGTGGCGAGCGCTGCGCGCGAGTGAAGCAATGGG 793
QY 240 YSerArgThrArgasp 245
Db 794 CAGCGGAAACCGCGA 809
RESULT 11
BG829472 845 bp mRNA linear EST 22-MAY-2001
LOCUS BG829472
DEFINITION Homo sapiens CDNA clone IMAGE:489112 5',
mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheraia; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 845)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1790 row: 9 column: 17
High quality sequence stop: 829.
Location/Qualifiers
1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:489112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;

	LOCUS	DEFINITION	ACCESSION VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE AUTHORS TITLE JOURNAL COMMENT
Oy	61	LeusSerGIugInValGIngIUGlUleueuSerSerGInValThrGIingIuLeuArGaIA	80				
Dd	253	CTGTCTGAAGCAGAGTCCAGAGAGAGCGTGCCTCACCTCCAGATGCCAGAACTGAGGGCG	312				
Oy	81	LeuMetAspGIUThrMetIysGIuLeuLysAlaTYrLYSserGIuLeuGIuIngIuLeu	100				
Dd	313	CTGATGGACAGAACCATGTAAGAGATTGAAGCGCTCAAAATTCGAACTGGAGAACACTG	372				
Oy	101	ThrProValAlaGIuGIuThrTrArgAlaArgLeuSerLYSGluLeuGIuAlaGIuAla	120				
Dd	373	ACCcgggtgcgagacagacgcggccacgcgctgtgtccaagacccttcaggcgccagagcc	432				
Oy	121	ArgLeuGIuAlaAspMetGIuAspValCYSGIYArGLeuValGIuThrArgGIuVal	140				
Dd	433	CcgcTggccgccgaatgtagagacctgtgcggccccccttggttcaattaccgcggcagagtg	492				
Oy	141	GlnAlaMetLeuGIuGIuInsErThrGIuGIuLeuArGVAlArgLeuAlaserHisLeuArg	160				
Dd	493	CAGGCAATGCTCGGCACAAGACACCGAAGAGCTCGCGGgtgcgccttcgccctccactgcgcc	552				
Oy	161	LysLeuArGVAlnArGLeuLeuArGVAspAlaAspAspleuGIuLysArgLeuAlaValTYr	180				
Dd	553	AAGCTGCGTTAAGCGGCTCTCCCGCATGCGCATGACCTGCAGAAGCCGCTGCAGATGATC	612				
Oy	181	GlnAlaGIuAlaArGVGIuGIuValaGIuArgLYLeuSerAlaIIlaArgLIuArgLeuGIy	200				
Dd	613	CAGGCGGGGGCCCCGGAAGGGCCCGCAGGCGGCTGTCCACCGCCATCCGAGCGCTTGggg	672				
Oy	201	ProLeuValGIuGIuGIuArgValArGVAlaArGVAlaThrValGIYSerLeuAlaGIuInPr	220				
Dd	673	CCCCTGGTGAACAAGGGCGCGGTGCGGGCCCGCACTGTGTGGCTCCCTGGCCAGCCAGCGG	732				
Oy	221	LeuGIuGIuAlaArGVAlaGIuAlaTrpGIuGIuArgLeuArGVAlaArgMetGIuMet	239				
Dd	733	CTACAGGAGCGGGCCAGGCGCTGTGGGGCGAGCGGTG-CGCGGCGGATGATGAGAGATG	788				
RESULT 13	B1551475	919 bp mRNA linear EST 05-SEP-2001					
LOCUS	B1551475	60319431.F1 NIH_MGC_95 Homo sapiens CDNA clone IMAGE:5274003 5'					
DEFINITION	mRNA sequence.						
ACCESSION	B1551475						
VERSION	B1551475.1	GI:15438787					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
TITLE	1 (bases 1 to 919)						
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/						
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nhl.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshlyuki and Piero Carninci (RIKEN) DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Gene distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLMK11691 row: 1 column: 04 High quality sequence stop: 812.						

```

BASE COUNT      172 a      270 c      363 g      114 t
ORIGIN
Alignment Scores:
Pred. NO.:      5.17e-110      Length:      919
Score:          1145.00      Matches:      246
Percent Similarity: 93.63%      Conservative: 4
Best local Similarity: 92.13%      Mismatches: 16
Query Match:     72.24%      Indels:      4
DB:              13      Gaps:      0
US-09-827-854-19 (1-317) x B1551475 (1-919)

```

BASE COUNT	172 a	270 c	363 g	114 t
ORIGIN				

Alignment Scores:

Pred. No.:	1.17e-110	length:	919
Score:	1145.00	Matches:	246
Percent Similarity:	93.63%	Conservative:	4
Best local Similarity:	92.13%	Mismatches:	16
Query Match:	72.24%	Indels:	4
DB:	13	Gaps:	0

US-09-827-854-19 (1-317) x BI551475 (1-919)

QY	1	MeysValleuTPPAlaIalaIeuleValThrPheuleuIaGlyCysGlnAlaIaVal	20
Dp	73	ATGAAGTTCGTGTGGGCGCTTCTGTCATTCCTGGAGGATGCCAGGCCAAAGTG	133
QY	21	GlueIuAlaValGIuThrGIuProGIuPheGluuArrGlnIuThrGIuTrpGlnSer	40
Dp	133	GAGCAAGCGGTGGAGACAGGCCGAGACCCGAGCTGGCGCAGCAGACCGAGTGGCAAGC	192
QY	41	GlyGlnArGTrpGIuLeuAlaIaLeuGlyIaRphetrPaspIyrieuArGTrpValGIuThr	60
Dp	193	GGCCAGCGCTGGGAACCTGGCACTGGGTGCTTTTGGGATTAACCTGGCTGGGTCAACA	252
QY	61	LeuSerGIuGIuValGIuGlnGIuLeuLeuSerGIuValThrGIuGlnIuLeuArAla	80
Dp	253	CTGTCTGACGAGGTGCAGAGAGAGCTGCTCACTGCCAGGTCCACGAACTAGAGCG	312
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaIyTrpLysSerGIuLeuGIuGlnIuLeu	100
Dp	313	CTGATGGACGAGACCATTAAGAGATTGAAGCCATCAAACTCGCAACTGGAGAACTAG	372
QY	101	ThrProValAlaGIuGIuThrArAlaIaArgLeuSerLysGIuLeuGlnAlaIaGlnAla	120
Dp	373	ACCCGGTGGCGGAGAGAGACGGGGCGCTGTCCAAAGAGCTGACAGCGCGCCAGGCC	432
QY	121	ArgLeuAlaIaAspMetGIuAspValCysGIuArgLeuValGlnTrpArgGlyGIuVal	140
Dp	433	CGGTGGCGCGGACATGAGAGAGTGTGGCGCGCTGTGTGCATGCCCGGGAGGTG	492
QY	141	GlnAlaMetLeuGlnGIuInsTrThGIuGIuLeuArGValArgLeuAlaSerHisLeuArg	160
Dp	493	CAGGCCATGCTCGGCCAGAGCACCAGAGAGCTGGGGTGGCCCTCGCTCCACCTGGC	553
QY	161	LysLeuArGTrpGIuArgLeuLeuArGAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Dp	553	AAGTCGCTAAGCGGCTCTCTCCGCGAGCCGATGATCTGCAGAAAGCCCTGCGAGTAC	612
QY	181	GlnAlaGlyAlaArGGIuGlyIaGlnIaGlnArgGlyLeuSerAlaIleArGValArgLeuGly	200
Dp	613	CAGCGCGGGGCCCGCATGTGGCGCGAGCGGTGCTGACGCCCATCCCGAGCGCTGGG	672
QY	201	ProLeuValGIuGlnGIuArgValArgAlaAlaThrValGlySerLeuAlaGlyGIuPro	220
Dp	673	CCCTGTGTGAACAGGGCCCGCTGCGGGCCGCCACTGTGGGCTTCCCTGGCGGACGCTT	732
QY	221	LeuGlnGIuArgAlaGlnAlaIaTrpGlyGIuArgLeuArGAlaArGMetGIuGIu-MetGI	240
Dp	733	A--CAGAGCGGGGCCAGGCCCTGGGGCGAGCGGnGCCGCGCGGAGTGAAGAAATGGG	790
QY	240	ySerArGTrpArGAspArgLeuAspGIuValLysGIuGlnValAlaGIuValArgAlaIaIa	260
Dp	791	CAGCGGACCCCGGACG--CTGGACGAAGTAAGAGAGCATGGCGGGGAGGG--TGCCGCAA	847

pbuescript KS-); Site_1: BamHI; site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 260 c 358 g 117 t 1 others

Alignment Scores:

Pred. No.:	5,69e-108	Length:	907
Score:	1125.50	Matches:	243
Percent Similarity:	94.62%	Conservative:	3
Best Local Similarity:	93.46%	Mismatches:	10
Query Match:	71.01%	Indels:	6
DB:	12	Gaps:	2

US-09-827-854-19 (1-317) x BG706129 (1-907)

```
OY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 75 ATGAGGTTCTGTGGGCTGCTGCTGTCACATTCTGTCAGAGATGCCAGCCAAAGTG 134
OY 21 GluGlnAlaValAlaGluThrLupProGluLeuArgGlnGlnThrLupLupSer 40
DB 135 GAGCAAGCGGTGGAGACAGCCGAGCCGAGCTCCGACAGACAGCCAGTGCCAGAGC 194
OY 41 GlyGlnAlaGlyTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 195 GGCACAGCGCTGGGAACCTGCTGCTTTGGATTACCTGGCTGGTGTCAGACA 254
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 255 CTGTCTGACAGGTGACGAGGAGCTGCTCAGCTCCAGTCCAGCAGCACTGAGGGCG 314
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 315 CTGATGGAGACGACCATGAGAGAGTTGAAGGCTTACAAATCCGAAGTGGAGACACTG 374
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
DB 375 ACCCGGTGGCGAGAGACGCGGCGACGCTGTCAGAGAGCTGACAGCGCGCAGGC 434
OY 120 ArgGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 435 CCGGCTGGGCGGACATGAGAGACGTGTGCGCGCTGTGTGCACTACCGCGCAGGT 494
OY 140 GlnAlaMetLeuGlnLysThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 495 GCAGGCGCATGCTGGCGACAGACGACGAGAGAGTGGGGTGGCCCTCCCTCCACCTGG 554
OY 160 GlySerLeuArgGlnArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 555 CAAGCTGGGTAAAGCGGCTCTCCGATGCCATGACCTGCAGAAAGCGCTGGCAGTGA 614
OY 180 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeu-G 200
DB 615 CCAGGCGCGGGGCGCGAGGCGCGGAGGCGGCTCAAGCGCCATCCGAGCGGCGCTGGG 674
OY 200 LysProLeuValGluGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyLup 220
DB 675 GCCCTTGGTGAACAGGCGCGCTGCGGCGCCACTGTGGCTCCCTGCGNCGCCAGC 734
OY 220 LysGluGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetG 240
DB 735 CGCTACAGAGCGGCGCAGGCTG---GGGCGAGCGGTGCGCGCGGATGAGAGATGG 791
OY 240 LysArgTrpThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 258
DB 792 GCAG-CGGACCCGG--ACGCTGAGAGATG-AAAGAGCAGTGGGGAAGTGTGCGC 842
```

Search completed: March 14, 2003, 20:14:18
Job time : 1293.69 secs

FT	CDS	1..954	/*tag= a	
FT		/product= "Human APOE protein"		
FT	variation	replace (13, C)		
FT		/*tag= a		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (31, G)		
FT		/*tag= b		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (364, A)		
FT		/*tag= c		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (388, C)		
FT		/*tag= d		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (498, G)		
FT		/*tag= e		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (526, T)		
FT		/*tag= f		
FT	variation	"Single nucleotide polymorphism (SNP)"		
FT		replace (622, A)		
FT		/*tag= g		
FT		/standard_name= "Single nucleotide polymorphism (SNP)"		
PN	WO200179234-A2.			
XX				
PD	25-OCT-2001.			
XX				
PR	16-APR-2001; 2001WO-US12303.			
XX				
PR	14-APR-2000; 2000US-197188P.			
XX				
PA	(GENA-) GENAISSANCE PHARM INC.			
XX				
PI	Choi JY, Kilem SE, Koshy B, Lee HH;			
XX				
DR	WPI: 2002-075064/10.			
XX	P-PSDB; AAE15158.			
PT	Genotyping human apolipoprotein gene of individual for determining			
PT	haplotype of individual, involves determining identity of nucleotide			
PT	pair at specific polymorphic sites for two copies of gene -			
XX				
PS	Claim 26; Fig 2; 78pp; English.			
XX				
CC	The patent discloses novel genetic variants of human apolipoprotein			
CC	E (APOE) gene. The invention also relates to compositions and methods			
CC	for haplotyping and/or genotyping the APOE gene. The haplotyping			
CC	methods of the invention are useful for improving the efficacy and			
CC	reliability of several steps in the discovery and development of			
CC	drugs for treating diseases associated with APOE activity, e.g.			
CC	familial dysbetalipoproteinemia, type III hyperlipoproteinemia,			
CC	atherosclerosis, and Alzheimer's disease. They are useful to validate			
CC	APOE as a candidate agent for treating a specific condition or diseases			
CC	predicted to be associated with APOE activity and in the design of			
CC	clinical trials of candidate drugs for treating a specific condition			
CC	or disease predicted to be associated with APOE activity. Genotyping			
CC	or haplotyping methods are useful to screen for compounds targeting			
CC	APOE to treat a specific condition or disease associated with APOE			
CC	activity. The present sequence is a cDNA encoding human APOE protein.			
XX				
XX	APOE gene is located on chromosome 19q13.2.			
SO	Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;			
Alignment Scores:				
Pred. No.:	1.02e-114	Length:	954	
Score:	1585.00	Matches:	317	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	24	Gaps:	0	

[illegible]

XX Human apolipoprotein-E.
 DE
 XX Apolipoprotein-E; hyperlipidemia; arteriosclerosis. ss.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 15..968
 FT /*tag= a
 XX
 XX AU8547513-A.
 XX
 XX 24-APR-1986.
 XX
 XX 17-SEP-1985; 85AU-0047513.
 XX
 XX 11-JUN-1985; 85JP-0126989.
 XX
 XX (MITU) MITSUBISHI CHEM IND KK.
 XX
 DR WPI: 1986-150217/24.
 DR P-PSDB: AAB60507.
 XX
 PT New DNA sequence coding for human apolipoprotein-E - and
 PT expression vectors and transformed cells contg. it
 XX
 XX Disclosure: Fig 2; 45pp; English.
 XX
 CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.
 XX
 SQ Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other:
 Alignment Scores:
 Pred. No.: 1,21e-114 Length: 1110
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0
 US-09-827-854-15 (1-317) x AAN60409 (1-1110)

DB 375 CGCTGGGCGCCGACATGGAGGACGTGTGCGGCCCTGTGTCAGTACCGCGGAGGTG 434
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCCATGCTCTGGCCAGAGCACCGAGAGAGCTGCGGGTCCGCTGCCCTCCACCTGGCC 494
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 495 AAGCTGCGTAAGCGGCTCTCCGCGCATGTGACACTGACAGAGCGGCTGGCAGTGTAC 554
 QY 181 GlnAlaGlyAlaArgGlyGluValGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 555 CAGGCCGGGGCCCGGAGGGCCCGCAGCGCGGCTCTAGCGCATCCGCGAGCGCCCTGGGG 614
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGlyInPro 220
 DB 615 CCCCTGTGTGAGACAGGCGCGGTGGGGCCGCACTGTGTGCTCTGACCGGCCAGCGG 674
 QY 221 LeuGlnGluArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGluLeuMetGly 240
 DB 675 CTACAGAGAGCGGGCCCGCAGGCTGTGGGGCGAGCGGTGGCGCGGATGAGAGATGGGC 734
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 DB 735 ACCCGGACCCCGGACCGGCTGTGAGCAGGTGAAGAGAGAGTGGCGAGGTGGCGCCCAAG 794
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
 DB 795 CTGGAGAGACCGCCAGACATACGCTGACAGCGCGAGGCTTCACAGGCCCGCTCAAG 854
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 DB 855 ACCTGGTGTGACCCCTGTGTGAACAGACATGACGCCGCAAGTGGCGGCTGTGTGAAGAAG 914
 QY 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
 DB 915 GTGCAGGCTGCGCGTGGGACCCAGCCCGCCCTGTGCCAGCAGACATCAC 965
 RESULT 3
 ID ABA83113 standard; DNA; 1147 BP.
 XX ABA83113;
 AC ABA83113;
 AC ABA83113;
 AC ABA83113;
 DT 08-FEB-2002 (first entry)
 DE Apolipoprotein E ovarian tumour marker gene. SEQ ID NO:63.
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytostatic;
 KW gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 OS
 XX
 XX WO20015177-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 03-APR-2001; 2001WO-US10947.
 XX
 XX 03-APR-2000; 2000US-194336P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 PI

US-09-827-854-15 (1-317) x AAD22052 (1-1156)

QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGTTCTGTGGGCTGCTGTGTCACATTCTGGCAGATGCGCAGAGCTG 120
 QY 21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGACAGCCGAGCTGCGCCAGCAGACAGGAGGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 Db 181 GGCCAGCGCTGGAGACTGGACACTGGTGCTTTGGGATTACTGCTGGCTGGGTGACAGCA 240
 QY 61 LeuSerGluGlnValAlaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 241 CTGTCTGAGCAGGTGACAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGGAACTGAGGCGC 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
 Db 301 CTGATGAGCAGACCATGATGAGAGTGTGAAGGCTTCAAAATGGAAGTGGAGAACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 361 ACCCGGTGGGCGAGAGAGCGGGCAGCGCTGTCCAGAGAGCTCAGAGCGCGCAGGCC 420
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
 Db 421 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGCCCTGCTGTCAGTACCGGCGGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 481 CAGGCCATGCTGGCCAGAGACCCAGAGAGTCTGGGTGGCTGCTGCCACTGGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 Db 541 AAGCTGCGCAGCGGCTCTCCGCGATGCGCATGACGTGACAGAACGCGCTGGCACTGAC 600
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
 Db 601 CAGGCGGGGGCGCGAGGGCGCCGAGCGGCTGACCGCATCGGAGCGGCGCTGGGG 660
 QY 201 ProLeuValAlaGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db 661 CCGCTGGTGGAAACAGGGCGCGCTGGGGCGGCACTGTGGGTCTCTGGCGGCCAGCCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysLysLeuArgAlaArgMetGluGluMetGly 240
 Db 721 CTACAGGAGCGGCGCCAGGCGCTGGGGCGAGCGGCTGCGCGGATGAGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db 781 AGCCGAGCGCGGAGCGCTGGAGCAGGTGAAGGAGAGGTGGCGAGAGTGGCGCCCAAG 840
 QY 261 LeuGluGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 Db 841 CTGGAGGAGAGAGCGCCAGCAGTACGCTGCGAGCGCGAGGCTTCCAGGCCCGGCTTCAAG 900
 QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
 Db 901 AGCTGTTGAGCGCCCTGGTGGAAACATGACGCGCCAGGTGGGCGGCTGGTGAAGAAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 Db 961 GTGAGGCTGGCGGAGCAGCGCGCCCTGTGCCAGGAGCAATCAC 1011

RESULT 7
 AAF84314
 ID AAF84314 standard; CDNA; 1156 BP.
 AC AAF84314;
 XX
 XX

DT 21-JUN-2001 (first entry)
 XX Human ApoE2 coding sequence.
 DE Human ApoE2 coding sequence.
 XX Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
 KW Homo sapiens.
 OS Homo sapiens.
 XX Key location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human ApoE2"
 PN JP2001017028-A.
 PD 23-JAN-2001.
 XX 28-APR-2000; 2000JP-0128919.
 PF 06-MAY-1999; 99JP-0125647.
 PR (MITU) MITSUBISHI CHEM CORP.
 PA WPI: 2001-285406/30.
 DR P-PSDB; AAB80996.
 XX New apoE humanized mammalian cell useful for screening for agents
 PT useful for treating or preventing Alzheimer's disease and
 PT arteriosclerosis -
 PS Disclosure; Page 11-12; 22pp; Japanese.
 XX The present invention relates to an ApoE humanised mammalian cell. The
 CC present sequence is the coding sequence for human ApoE2, which was used
 CC in the method of the present invention. The ApoE humanised mammalian cell
 CC can be used for screening for agents useful for treating or preventing
 CC Alzheimer's disease and arteriosclerosis.
 XX Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
 SO Alignment Scores:
 Pred. No.: 5.3e-114 Length: 1156
 Score: 1577.00 Matches: 316
 Percent Similarity: 99.68% Conservative: 0
 Best Local Similarity: 99.68% Mismatches: 1
 Query Match: 99.50% Indels: 0
 Gaps: 0
 US-09-827-854-15 (1-317) x AAF84314 (1-1156)
 QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 61 ATGAAGTTCTGTGGGCTGCTGTGTCACATTCTGGCAGATGCGCAGAGCTG 120
 QY 21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db 121 GAGCAAGCGGTGGAGACAGACAGCCGAGCTGCGCCAGCAGACAGGAGGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 Db 181 GGCCAGCGCTGGAGACTGGACACTGGTGCTTTGGGATTACTGCTGGCTGGGTGACAGCA 240
 QY 61 LeuSerGluGlnValAlaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 241 CTGTCTGAGCAGGTGACAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGGAACTGAGGCGC 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
 Db 301 CTGATGAGCAGACCATGATGAGAGTGTGAAGGCTTCAAAATGGAAGTGGAGAACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 361 ACCCGGTGGGCGAGAGAGCGGGCAGCGCTGTCCAGAGAGCTCAGAGCGCGCAGGCC 420

```
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlyVal 140
    |||||||
Db 421 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGCGCTGTGACGATCCCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 481 CAGGCCATGCTCGGCCAGACACCGAGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||||
Db 541 AAGCTGCTGAAGGCGGCTCCCGCGATGCCGATGACCTGAGAAAGTCCCTGGCAGTTCAC 600
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||||
Db 601 CAGGCCGCGGCGCGGAGGCGCGGAGCGCGGCTCAGCCATCCCGCAGACCGCTGGGG 660
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
    |||||||
Db 661 CCGCTGTGGACAGAGGCGCGCTGGCGGCGCGCACCTGTGGCTCCGCGCGCGCACGCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTyrGlyGluArgLeuArgAlaArgMetGlnGluMetGly 240
    |||||||
Db 721 CTACAGAGAGCGGCGCGGCTGGGGGAGCGGCTGGCGCGCGGAGTGAAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
    |||||||
Db 781 AGCGCGGACCGCGACCGCTGAGAGAGTGAAGAGAGCGAGTGGCGGAGTGGCGCGCACG 840
QY 261 LeuGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
    |||||||
Db 841 CTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
    |||||||
Db 901 AGCTGTGTGAGAGCGCTCGTGGTGAAGACATGACAGCGCGAGTGGCGGCTGTGTGAAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
    |||||||
Db 961 GTCCAGAGCTCCGCTGGGACACAGCGCGCGCTGTGCCAGCAATCATC 1011
RESULT 8
ID AAD22049 standard; DNA; 1156 BP.
AC AAD22049;
XX
DT 12-FEB-2002 (first entry)
DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
KM Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT sig_peptide 61..114 /product= "Human apoE isoprotein, apoE2"
FT mat_peptide 115..1011 /tag= a
FT sig_peptide 61..114 /tag= b
FT mat_peptide 115..1011 /tag= c
FT /product= "Mature human apoE isoprotein, apoE2"
XX
PN WO200177136-A1.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001MO-US11358.
XX
PR 06-APR-2000; 2000US-0544386.
```

```
PR 04-OCT-2000; 2000US-0679088.
PR 05-APR-2001; 2001US-0827854.
XX
XX (KOSP-) KOS PHARM INC.
PA (UYBO-) UNIV BOSTON.
XX
PI Zannis VI, Kypros KE;
XX
DR WPI; 2002-010885/01.
DR P-PSDB; AAE13295.
XX
PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT cholesterol, delaying the onset of or treating atherosclerosis in
PT mammal, without inducing hypertriglyceridaemia.
XX
PS Claim 14; Page 81-82; 91pp; English.
XX
CC The present sequence is a human apolipoprotein E (apoE)
CC isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC lowering cholesterol, delaying the onset of atherosclerosis,
CC treating or regressing atherosclerosis without inducing
CC hypertriglyceridaemia, in a mammal lacking an endogenous,
CC normally functioning apoE gene or low density lipoprotein (LDL)
CC receptor or is at risk for developing atherosclerosis due to
CC accumulation of lipoprotein remnants in the bloodstream or having
CC a defect in remnant removal.
XX
SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
Alignment Scores:
Pred. No.: 5,3e-114 Length: 1156
Score: 1577.00 Matches: 316
Percent Similarity: 99.68% Conservative: 0
Best Local Similarity: 99.68% Mismatches: 1
Query Match: 99.50% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-15 (1-317) x AAD22049 (1-1156)
QY 1 MetLysValLeuTrpAlaAlaLeuValAlaThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||||
Db 61 ATCAAGGTTCTGTGGGCTGCGTGTGTCACATTCTCGCAGAGATGCCAGGCCAAGGTG 120
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnTrpGlnSer 40
    |||||||
Db 121 GAGCAAGCGGTGGAGAGACAGAGCGCGAGCGCGAGCTCGCGCAGAGACCGAGTGGCAGAC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||||||
Db 181 GGCACAGCGGTGGAGACTGACAGCGGCTCTTTGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
    |||||||
Db 241 CTGTCTGAGCAGGTGCAGAGAGAGAGCTGTGCTCAGCTCCAGAGTACCCAGAACTAGAGCGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnGlnLeu 100
    |||||||
Db 301 CTGATGAGACAGACATGAAGAGGTGAAGGCTTACAAATCGGAATCGAGAGAACTG 360
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnGlnGlnAlaGlnAla 120
    |||||||
Db 361 ACCCGGCTGGGAGAGAGAGCGCGGAGCGCTGTCCAAAGAGCTGAGAGCGCGCGCGCGCG 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGlyVal 140
    |||||||
Db 421 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGCGCTGTGACGATCCCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||||
Db 481 CAGGCCATGCTCGGCCAGACACCGAGAGAGCTGGGGTGGCTCGCTCCGCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||||
```


Db	721	CTACGAGAGCGGGCCAGCGCTGGGGCGAGCGGCTGGCGCGCGCGGATGAGAGATGGGC	780		
Qy	241	SeRaGtrhrArGAsPaRgLeuAsPgluValLySgluGluGlnValAlaGluValArGAlaLys	260		
Db	781	AGCGGACCCCGACCGCGCTGGAGAGAGGTGAAGAGACAGAGTGGCGGAGGTGGCGCCAC	840		
Qy	261	LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys	280		
Db	841	CTGGAGGAGCGAGCGCCACAGATACGCTGCGAGCGCCGAGGCGCTTCCAGGCGCCGCTCAC	900		
Qy	281	SeRTPpHeGluPProLeuValGluAspMeGlnArgGlnIrrPrAlaGlyLeuValGluLys	300		
Db	901	AGCTGTTTCAGACCCCTGTGTGAAGACTGAGAGGCCAGTGGGCGGGGCTGTGTGAGAA	960		
Qy	301	ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis	317		
Db	961	GTGCGAGCTGCCGTGGGACACAGCGCCGCCCTGTGCCGACGACATTCAC	1011		
RESULT 10					
ID	AA06957	AA06957 standard; cDNA to mRNA; 1157 BP.			
XX	AA06957;				
AC	AA06957;				
DT	19-JUN-1996	(first entry)			
XX					
DE	Human apolipoprotein-E (ApoE) cDNA.				
XX					
KM	Recombinant; human; apolipoprotein-E; ApoE; insect cells; larva;				
KW	Manduca sexta; Autographica californica californica nuclear polyhedrosis virus;				
RV	haemolymph; lipid complex; biologically active; ds.				
XX					
OS	Homo sapiens.				
XX					
EH	Key	Location/Qualifiers			
FT	Misc_feature	62..119			
FT		/*tag= a			
FT		note= "misc_signal"			
FT	CDS	62..1015			
FT		/*tag= b			
XX					
PN	US5472858-A.				
XX					
PD	05-DEC-1995.				
XX					
PE	04-JUN-1991;	91US-0709949.			
XX					
PR	04-JUN-1991;	91US-0709949.			
XX					
PA	(WISC) WISCONSIN ALUMNI RES FOUND.				
PI					
PI	Attie AD, Beckage NE, Gretch DG, Sturley SL;				
XX					
DR	WPI; 1996-029812/03.				
DR	P-PSDB; AAR86791.				
XX					
PT	Prodn. of recombinant apolipoprotein E in insects - by infecting				
PT	Manduca sexta larvae with recombinant Autographica californica				
XX	nuclear polyhedrosis baculovirus vector.				
XX					
Disclosure; Columns 11-14; 10pp; English.					
CC	Recombinant human apolipoprotein-E (ApoE) (AAR86791) can be produced				
CC	by preparing a genetic construct (contg. an ApoE-encoding sequence,				
CC	e.g. AAT06957, and flanking regulatory sequences enabling the protein				
CC	to be expressed in insect cells), which is then introduced into a				
CC	Manduca sexta larva (using a recombinant Autographica californica				
CC	nuclear polyhedrosis virus) and recovering the protein from the				
CC	haemolymph of the larval host. The ApoE produced is in a form				
CC	sufficiently complexed with lipids to be biologically active, which				
CC	cannot be achieved in insect cell cultures, and can therefore be				
CC	used in therapeutic applications.				

```

XX      Sequence 1157 BP: 212 A; 370 C; 426 G; 149 T; 0 other:
SQ
Alignment Scores:
Pred. No.:      6,34e-114      Length:      1157
Score:          1576.00        Matches:      315
Percent Similarity: 99.37%      Conservative: 0
Best Local Similarity: 99.37%      Mismatches:  2
Query Match:     99.43%         Indels:      0
DB:              17            Gaps:        0

```

US-09-827-854-15 (1-317) x AAT06957 (1-1157)

QY	1	MellysValLeuTPPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	62	ATGAAGGTTCTGTGGGCTGCTGTCTGTGTCAATTCCTGGAGAGATGCCAGGCCAAGGTG	121
QY	21	GIuGlnAlaValAGluThrGluProGluuProGluLeuAArgGlnGlnThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGGAGACACAGAACCCGAGCTGCCGCCAGCAGACCGAAGTGGCAGACC	181
QY	41	GlyGlnAArgTPGluLeuAlaLeuGluYArpPheTPAsPTyLeuAArgTPValGlnThr	60
Db	182	GGCCAGCGCTGGGAAGTGGCACTGGGTGGCTTTTGGGATTCACGTGGCTGGGTCAACA	244
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuAArgAla	80
Db	242	CTGTCTGTGAGCGAGGTGCACAGAGAGACTGTCTGACGTCCCAAGTACCCCAAGAACTAGAGGCG	303
QY	81	LeuMeTAspGluTPuMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGluLeu	100
Db	302	CTGATGGACGAGACCATTAAGATTTGAAGGCTTACAAATCGGAACATGGAGAAACAATCG	363
QY	101	ThrProValAlaGluGluThrArgAlaAArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	362	ACCCGGAGTGGGAGGAGAACCCGGGCAAGCTGTCCAAAGAGCTGCAGACGCGCCAGGCC	421
QY	121	ArgLeuGlnYAlaAspMetGluAspValCysGlyLysLeuValGlnTyrArgGlyGluVal	140
Db	422	CGGCTGGGGCGCGACATGGAGACGTGTGGCGCCGCTGTTGGCATACCGCGGGCAGGTG	481
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuAArgValAArgLeuAlaSerHisLeuAArg	160
Db	482	CAGGCAATGCTCGGGCCAGACACCGAGAGAGTGGGGTGGCCCTCCGCTCCACCTCCGCG	541
QY	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTyr	180
Db	542	AAGCTGCGTTAAGCCGGCTCTCCGGATCCCATATACCTGCGAAGGCCCTGGCAGTTAC	601
QY	181	GlnAlaGluYAlaAArgGluGluYAlaGluAArgGlyLeuSerAlaAlaLeuAArgLysGluY	200
Db	602	CAGGCGGGGGCCCGAGAGGGCGCGGAGCGGCTCAAGCCGCAATCCCGAGCCGCTGGGG	661
QY	201	ProLeuValGluGlnGlyAArgValAArgAlaAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	662	CCCTGTGGTGAACAGGGCGCGGTGGGAGCCCACTGTGGGCTCTCCGTGGCGGCACCGG	721
QY	221	LeuGlnGluAArgAlaGlnAlaAArgGlyGluAArgLeuAArgAlaAArgMetGluMetGly	240
Db	722	CTTACAGAGACGGGCCCAAGCCTGGGGGAGAGCGCTGGCCCGCGCATGGAGAGATGGCC	781
QY	241	SerAArgThrAArgAspAArgLeuAspGluValLysGluGlnValAlaGluValAArgAlaLys	260
Db	782	AGTGGGAGCCCGACCCGCTGGAGAGAGTGAAGAGCAGGTGGCGGAGTGGCGGCCAAG	841
QY	261	LeuGluGluGlnAlaGlnGlnIleAArgLeuGlnAlaGlnAlaPheGlnAlaAArgLys	280
Db	842	CTTGAGAGACAGGCCCAAGCAATACGCTCGAGGCCCAAGGCTTCCAGGCCCGCTCAAG	901
QY	281	SerTPuPheGluProLeuValGluAspMetGlnAArgIleTPuAlaGlyLeuValGluLys	300
Db	902	AGCTGGTTCCGAGCCCTGGTGAAGACATGACGCCAGTGGCGGGCTGTGTGAAG	961

OY 301 ValGlnAlaIaValGlyThrSerAlaAlaProSerAspAsnHis 317
|||||
Db 962 GTGCAAGGCTGCGGTGGGACACAGCCGCCCTGTGTGCCAGGACATTCAC 1012

RESULT 11
ABN95746
ID ABN95746 standard; DNA; 1157 BP.
XX
AC ABN95746;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2244 used to diagnose liver cancer.
XX
KM Gene: liver cancer; ds: hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumour; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN MO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001MO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-De-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX
PS Claim 1; SEQ ID NO 2244; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 6,34e-114 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
Gaps: 0
DB: 24

US-09-827-854-15 (1-317) x ABN95746 (1-1157)

OY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 62 ATGAAGGTTCTGTGGGCTGCGTTCACATTCTCTGCGAGATGCCAGGCCAAGGTG 121

OY 21 GlnAlaIaValAlaGlyThrGluProGluProGluLeuArgGlnGlnThrGluThrPheGlnSer 40
|||||
Db 122 GAGCAAGCGGTGGAGACAGAGCCGAGCGCTGCGGCAGACAGACCGAGTGGCAGAC 181

OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
|||||
Db 182 GGCACAGCGCTGGAACTGGACACTGGGTGCTTTTGGATTACTCGCTGGGTGACAGACA 241

OY 61 LeuSerGluGlnValAlaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 242 CTGTCTGAGCAGTGCAGAGAGAGCTGTCCACTCCCAAGTCCACCAAGAACTGAGGCGG 301

OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||
Db 302 CTGATGAGACGAGACCATGAGAGAGTTGAAAGCTTCAAAATCCGAATCGAGGAGACACTG 361

OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 362 ACCCGGTAGCGAGAGAGAGCGCGGCGACGCTGTCCAAAGAGACTGCAGACGCGCAGGCG 421

OY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
|||||
Db 422 CGGTGGGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCTGCGGCGAGGTG 481

OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 482 CAGGCCATGCTGGCGCAGAGACAGCAGAGAGCTGGCGGTGCGCTGCCCTCCACTGCGC 541

OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 542 AAGCTGCGTAAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGCGACGTGAC 601

OY 181 GlnAlaGlyAlaArgGluGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 602 CAGCGCGGGCGCGGAGGCGCCGAGCGCGCTTCAGCGCATCCGAGAGCGCTGTGGGG 661

OY 201 ProLeuValAlaGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
|||||
Db 662 CCCCTGTGTGAACAGAGCGCGCTGTGGCGCGCCACTGTGTGCTCCCTGCGCGCCAGCGG 721

OY 221 LeuGlnGluArgAlaGlnAlaIleThrPheGluArgLeuArgAlaArgMetGlnGluMetGly 240
|||||
Db 722 CTACAGAGAGCGGCGCCAGCGCTGTGGCGAGCGGTGCGCGCGGATGGAGAGATGGGC 781

OY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys 260
|||||
Db 782 AGTCGGACCGCGGACCGCTGTGGAGAGGTGAAGGAGCAGGTGCGGAGCTGCGCCCAAG 841

OY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
|||||
Db 842 CTGGAGAGAGAGCGCCAGCAGATACGCTGCGAGCGCGGAGGCTTCCAGGCGCGCTCAAG 901

OY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGluLys 300
|||||
Db 902 AGCTGTGTGAGCGCCCTGTGGTGAACATGACAGCCGCAATGGCGCGCTGTGTGAAGAAG 961

OY 301 ValGlnAlaIaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
|||||
Db 962 GTGCAAGGCTGCGGTGGGACACAGCCGCCCTGTGTGCCAGGACATTCAC 1012

RESULT 12
ABK64514
ID ABK64514 standard; DNA; 1157 BP.
XX
AC ABK64514;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human benign prostatic hyperplasia gene #409.
XX
KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS Homo sapiens.

XX WO200212440-A2.
PN
XX
PD 14-FEB-2002.
XX
PF 07-AUG-2001; 2001WO-US24708.
XX
PR 07-AUG-2000; 2000US-223323P.
PR 05-JUN-2001; 2001US-0873319.
XX
PA (GENE-) GENE LOGIC INC.
PA (NISB) JAPAN TOBACCO INC.
PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
DR WPI: 2002-257476/30.
XX
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
XX
XX
PS Disclosure: Page 239-240: 444pp: English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 1157 BP: 212 A; 370 C; 426 G; 149 T; 0 other:

Alignment Scores:
Pred. No.: 6.34e-114 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: 24 Gaps: 0

US-09-827-854-15 (1-317) x ABK64514 (1-1157)
QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAAGTTCTGTGGCTCGTGTGTCACATTCTCGGAGGAGCCAGGCCAAGG 121
QY 21 GUGlnAlaValAlaGlnThrGluProGluProGluLeuArgGlnGlnThrGlnPrgIns 40
DB 122 GAGCAAGCCGTGGAGACACAGCCGGAGCCGAGCTGCGCAGCAGCCAGTGGCAGAC 181
QY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCAGAGCCGTGGAACTGGCAGCTGGCTTTGGATTACTGCGCGGGGCGAGACA 241
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
|||||

DB 242 CTGTCTGAGCAGGTGCGAGGAGAGACTGCTAGCTCCCAAGTCACCCAGAACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
DB 302 CTGATGACGACAGACCATGAAAGAGTTGAAGCCCTACAAATCCGAATCGAGCAACTG 361
QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAla 120
DB 362 ACCCGGTGAGCGAGAGACGCGGCGACGCTGTCCAGAGAGCTGCAGACGCGCGAGGCC 421
QY 121 ArgLeuGlnAlaAspMetLysPvalLysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 422 CGGCTGGCGCGACATGAGGACGTGTGGCGCGCTGTGTGACATCCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnLysInsThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGCCATGCTCGGCCAGACAGACCGAGAGCTGCGGGTCCGCTCCCACTGCCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
DB 542 AAGCTCGTAAAGCGCTCTCCCGCATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaThrLeuArgLysGln 200
DB 602 CAGCGCGGGCCCGCAGAGGCGCGCGCGCTCAGCGCATCCGACGCGCTGGG 661
QY 201 ProLeuValGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
QY 221 LeuGlnGlnArgAlaGlnAlaThrPrgLysGlnArgLeuArgAlaArgMetGlnMetGly 240
DB 722 CTACAGACGCGGGCCAGCGCTGTGGCGAGCGGTGCGCGCGCGATGAGAGATGGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAlaLys 260
DB 782 AGTGGACCGCGACCGCTGTGAGGAGGTGAAGAGACAGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGlnGlnAlaGlnGlnAlaArgLeuGlnAlaAlaPheGlnAlaArgLeuLys 280
DB 842 CTGGAGACAGACGCGCAGAGATAGCGCTGACGCGAGCGAGCGCTTCAGGCCCGCTCAAG 901
QY 281 SerTrpPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
DB 902 AGTGTGTGTGAGCCCTGTGTGAGACATGACAGCGCAGTGGCGCGCTGTGTGAGAG 961
QY 301 ValGlnAlaAlaValGlnThrSerAlaAlaProValProSerAspAsnHis 317
DB 962 GTGCAAGCTGCGGTGGCACACAGCCGCCCTGTGCTCCAGGACATATCAC 1012

RESULT 13
ABL65450
ID ABL65450 standard; DNA: 1157 BP.
XX
AC ABL65450;
XX
XX 15-MAY-2002 (first entry)
DT
XX
XX Lung cancer related gene sequence SEQ ID NO:3787.
DE
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms's tumour; adenocarcinoma;
XX gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200194629-A2.
PN
XX
XX 13-DEC-2001.
PD
XX
XX 30-MAY-2001; 2001WO-US10838.
PF
XX

PR	05-JUN-2000;	2000US-209473P.	
PR	05-JUN-2000;	2000US-209531P.	
PR	18-SEP-2000;	2000US-233133P.	
PR	18-SEP-2000;	2000US-233617P.	
PR	20-SEP-2000;	2000US-234009P.	
PR	20-SEP-2000;	2000US-234034P.	
PR	20-SEP-2000;	2000US-234052P.	
PR	22-SEP-2000;	2000US-234509P.	
PR	22-SEP-2000;	2000US-234567P.	
PR	25-SEP-2000;	2000US-234923P.	
PR	25-SEP-2000;	2000US-234924P.	
PR	25-SEP-2000;	2000US-235077P.	
PR	25-SEP-2000;	2000US-235082P.	
PR	25-SEP-2000;	2000US-235134P.	
PR	25-SEP-2000;	2000US-235280P.	
PR	26-SEP-2000;	2000US-235637P.	
PR	26-SEP-2000;	2000US-235638P.	
PR	27-SEP-2000;	2000US-235711P.	
PR	27-SEP-2000;	2000US-235720P.	
PR	27-SEP-2000;	2000US-235840P.	
PR	27-SEP-2000;	2000US-235863P.	
PR	28-SEP-2000;	2000US-236028P.	
PR	28-SEP-2000;	2000US-236033P.	
PR	28-SEP-2000;	2000US-236034P.	
PR	28-SEP-2000;	2000US-236109P.	
PR	28-SEP-2000;	2000US-236111P.	
PR	29-SEP-2000;	2000US-236842P.	
PR	29-SEP-2000;	2000US-236894P.	
PR	02-OCT-2000;	2000US-237172P.	
PR	02-OCT-2000;	2000US-237173P.	
PR	02-OCT-2000;	2000US-237278P.	
PR	02-OCT-2000;	2000US-237294P.	
PR	02-OCT-2000;	2000US-237295P.	
PR	02-OCT-2000;	2000US-237316P.	
PR	03-OCT-2000;	2000US-237425P.	
PR	03-OCT-2000;	2000US-237598P.	
PR	03-OCT-2000;	2000US-237604P.	
PR	03-OCT-2000;	2000US-237606P.	
PR	03-OCT-2000;	2000US-237608P.	
PR	01-NOV-2000;	2000US-244867P.	
PR	01-NOV-2000;	2000US-245084P.	
XX	(AVAL-) AVALON PHARM.		
PA			
PI	Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
DR	WPI: 2002-188264/24.		
XX			
XX			
PT	Screening for anti-neoplastic agent involves exposing cells to a		
PT	chemical agent to be tested for anti-neoplastic activity, and		
PT	determining a change in expression of a gene of a signature gene set		
XX			
PS	Claim 1; SEQ ID 3787; 44pp; English.		
XX			
CC	The present invention describes a method (M1) for screening for an		
CC	anti-neoplastic agent. The method involves exposing cells to a chemical		
CC	agent to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I1) of a signature gene set, where (I1)		
CC	comprises a sequence (S) selected from 8447 sequences (given in AB61664		
CC	to AB170110), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (I1) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening		
CC	an anti-neoplastic agent, and can be used for producing a product which		
CC	is the data collected with respect to the anti-neoplastic agent as a		
CC	result of M1, and the data is sufficient to convey the chemical		
CC	structure and/or properties of the agent. M1 can be used in the		
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,		
CC	oesophageal, ovarian, kidney, prostate or pancreatic cancer,		
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer		
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine		
CC	carcinoma, papillary carcinoma and Wilm's tumour.		
CC			

QY 301 ValGlnAlaIaValGlyThrSerAlaIaProValProSerAspAsnHis 317
DB 962 GTGCAGAGCGCGCTGGGACACAGCGCCCGCTGTGCCAGCGACATCACC 1012

RESULT 14
AAN50450
ID AAN50450 standard; DNA; 1110 BP.
XX
AC AAN50450;
XX
DT 09-JAN-1992 (first entry)
XX
DE Sequence encoding human apolipoprotein E.
XX
KM Hyperlipaemia; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 15..968
FT /tag= a
FT mat_peptide 69..965
FT /tag= b

XX JP60118189-A.
XX
XX PD 25-JUN-1985.
XX
XX PF 29-NOV-1983; 83JP-0224980.
XX
XX PR 29-NOV-1983; 83JP-0224980.
XX
XX PA (MITU) MITSUBISHI CHEM IND KK.
XX
XX DR WPI: 1985-188003/31.
XX P-PDB: AAPS1204.
XX
XX PT DNA fragment - contg. DNA which codes human apolipoprotein E for
XX treatment of hyperlipaemia.
XX
XX PS Claim 3; Page 484; 8pp; Japanese.
XX
XX CC The sequence may be used to produce the apolipoprotein E, useful in the
XX treatment of hyperlipaemia.
XX
SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other;

Alignment Scores:
Pred. No.: 7,24e-114 Length: 1110
Score: 1575.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.37% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15 (1-317) x AAN50450 (1-1110)

QY 1 MetLysValLeuThrPalaIaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 15 ATGAAAGTTCTGTGGGCTGCTGCTGTCACATTCTCTGGCAGATGCCAGCGAAGTG 74

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
DB 75 GAGCAAGCGGTGGAGACAGAGCCGAGAGCTGCGCAGACAGACCGAGTGGCAGAGC 134

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 135 GGCACACCGCTGGGAACCTGCACCTGGGCTCTTTGGATTACCTGCGCTGGGTCAGACA 194

QY 61 LeuSerGluGluValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 195 CTGTCTGACACAGCTGCAGGAGGAGCTGCTGACCTCCAGAGTACCCAGGAAGTGAAGGCG 254

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
DB 255 CTGATGAGAGAGACCATGAAGAGTTGAAGGCTTACAAATCGAAGTGAAGCAACATG 314

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 315 ACCCGGTGGGAGAGAGAGCGGCGACGCTGTCCAAAGAGCTGAGCGCGCCAGGCC 374

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 375 CGGCTGGGCGGACATGAGAGACGTGTGCGGCGGCTGTGAGTGAACCGCGGAGGTG 434

QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 435 CAGGCCATGCTCGGCGCAGACACCGAGAGCTGCGGCTGCCCTCCCTCCACCTGCC 494

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaLysTyr 180
DB 495 AAGCTCGTAAAGGCGCTCTCCGATGCCGATGACCTCAGAAAGCGCTGGCAGTGTAC 554

QY 181 GlnAlaGlyAlaArgGluGluGluAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 555 CAGCGCGGGGCGCGAGGCGCGGAGCGGCGGCTCAGGCGCATCCGAGGCGCTGGG 614

QY 201 ProLeuValGluGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyInPro 220
DB 615 CCCCTGTGGAAACAGGCGCGCTGTGCGGCGCCACTGTGGCTCCCTGCGCGCGCCG 674

QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGluMetGly 240
DB 675 CTACAGAGAGCGGCGCCAGGCTGGGCGAGCGGCTCGCGCGCGGATGAGAGATGGGC 734

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 AGCGGAGCCCGGAGCGCGCTGTGAGCAGGTGAAGAGACAGTGGCGGAGCGCGCAAG 794

QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAlaArgLeuLys 280
DB 795 CTGGAGAGACAGGCGCCAGAGATACCCCTGCAGGCGGAGCGCTTCAGGCGCGCTCAAG 854

QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlyLys 300
DB 855 AGCTGTTTGAAGCCCTGTGTGAAGCATGCACGCCCACTGGGCGCGCTGTGGGAAG 914

QY 301 ValGlnAlaIaValGlyThrSerAlaIaProValProSerAspAsnHis 317
DB 915 GTGCAGAGCGCGCTGGGACACAGCGCCCGCTGTGCCAGCGACATCACC 965

RESULT 15
AAF84316
ID AAF84316 standard; CDNA; 1156 BP.
XX
AC AAF84316;
XX
XX DT 21-JUN-2001 (first entry)
XX
XX DE Human Apoe4 coding sequence.
XX
XX KM Human; Apoe4; Alzheimer's disease; arteriosclerosis; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human Apoe4"

XX JP2001017028-A.
XX
XX PD 23-JAN-2001.
XX
XX PF 28-APR-2000; 2000JP-0128919.
XX

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 34.7343 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALVTLFLGCGCAKV.....VEKVQAAVGTSAAPVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n_model -DEV=xlp
-Q=/cgn2.1/USPFO.spool/US09827854/rnatac_11032003_101611_27510/app_query.fasta.1.3576
-DB=Issued_Patents_NA -QPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=100 -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.cgn1.1.274 @runat.11032003_101611_27510 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEOURR -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARK_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCBUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1576	99.4	1157	1	US-07-709-949-1
2	1138	71.8	1126	4	US-08-949-155-5
3	1138	71.8	1126	4	US-09-819-964-5
4	1002	63.2	4267	4	US-08-949-155-51
5	1002	63.2	4267	4	US-09-819-964-51
6	984	62.1	660	2	US-07-726-306A-28
7	511	32.2	330	1	US-07-849-389-6
8	374	23.6	252	3	US-08-617-256-24
9	374	23.6	252	4	US-09-287-141-24
10	374	23.6	252	4	US-09-431-613-24
11	374	23.6	252	4	US-09-504-245-24
12	374	23.6	252	4	US-09-287-682-24

13	374	23.6	252	4	US-09-287-679-24	Sequence 24, Appl
14	374	23.6	252	4	US-09-397-766-24	Sequence 24, Appl
15	374	23.6	252	4	US-09-287-681-24	Sequence 24, Appl
16	374	23.6	252	4	US-09-495-444-24	Sequence 24, Appl
17	177.5	11.2	842	1	US-08-952-796-1	Sequence 1, Appl
18	174.5	11.0	801	4	US-07-959-946-4	Sequence 4, Appl
19	174.5	11.0	801	5	US-08-333-577-4	Sequence 4, Appl
20	174.5	11.0	801	5	PCT-US92-08634-4	Sequence 4, Appl
21	159	10.0	5661	4	US-08-938-105-2	Sequence 2, Appl
22	157.5	9.9	964	4	US-08-448-606-5	Sequence 5, Appl
23	154.5	9.7	863	1	US-08-448-606-7	Sequence 7, Appl
24	154	9.7	3256	2	US-08-968-751-3	Sequence 3, Appl
25	153	9.7	1879	4	US-09-750-580-2	Sequence 2, Appl
26	147.5	9.3	603	4	US-08-952-254-15	Sequence 14, Appl
27	146	9.2	8789	1	US-08-328-254-5	Sequence 5, Appl
28	146	9.2	10136	1	US-08-353-700-2	Sequence 2, Appl
29	146	9.2	10136	5	PCT-US95-16216-2	Sequence 2, Appl
30	145	9.1	4868	1	US-08-139-937-12	Sequence 12, Appl
31	145	9.1	4868	5	PCT-US93-11310-12	Sequence 12, Appl
32	145	9.1	6306	1	US-08-195-487-3	Sequence 3, Appl
33	145	9.1	6306	5	PCT-US93-06160-3	Sequence 3, Appl
34	144	9.1	1771	2	US-08-533-669A-7	Sequence 7, Appl
35	144	9.1	1771	2	US-08-511-872-1	Sequence 7, Appl
36	144	9.1	1771	4	US-09-183-861-7	Sequence 7, Appl
37	144	9.1	1771	4	US-09-022-765-7	Sequence 7, Appl
38	144	9.1	13121	4	US-08-961-527-126	Sequence 126, App
39	143.5	8.9	516	4	US-07-853-913-1	Sequence 1, Appl
40	141	8.9	516	4	US-09-183-861-34	Sequence 34, Appl
41	141	8.9	516	4	US-09-022-765-34	Sequence 34, Appl
42	141	8.9	6306	1	US-08-466-390-3	Sequence 3, Appl
43	141	8.9	6306	1	US-08-470-950-3	Sequence 3, Appl
44	141	8.9	6306	1	US-08-467-781-3	Sequence 3, Appl
45	141	8.9	6306	2	US-08-483-924-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: P. O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

[illegible][illegible]

```

QY 38 TTPGlnSerGlycInaArgTrpGluLeuAlaLeuGlyArgPheThrPaspTyrLeuArgTTP 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 TGGCAGGGAGCCAGCCCTGGGAGCAGGCCCTGGCGCTTCTGGATACCTGCGCTGG 218
QY 58 ValGlnThrLeuSerGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GTGCAAGTCCCTGTGTCAGCAAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 278
QY 78 LeuArgAlaLeuMetAspGluThrMetGlyGluLeuValAlaTyrGlySerGluLeuGlu 97
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CTGACGAGAGCTGATPAGAGGAGACATGAAAGAGTGAAGCCCTACCGGAGAGGAGTGGAG 338
QY 98 GluGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerGlyGluLeuGlnAla 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GCCAGAGCTGGGCCCTGGAGCCAGAGAGCCAGGCGCGCTGTCCAGAGAGCTGCAGAGCG 398
QY 118 AlAGlnAlaArgLeuGlnGlyAlaAspMetGlyAspValCysGlyArgLeuValGlnTyrArg 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 GCCAGAGCCCGCTGGGCGGAGCATGAGAGAGTGGCGCAACCGCTGTGTCTACCGC 458
QY 138 GlyGlnValGlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSer 157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 AGCAGAGTGCACAACATGTGTGGCCAGACACCGAGAGCTGGGAGCGCGCTGGCTTC 518
QY 158 HisLeuArgGlySerLeuArgGlyArgLeuLeuArgAlaAspAspLeuGlnGlyArgLeu 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 CACCTCGCAGAGCTGCGCAAGCGGCTGCTCCGCGACACCGAGCTGCAGAGCCCTG 578
QY 178 AlaValTyrGlnAlaGlnAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GCCGTGTACAGGCGGGGCGCTGGCGGAGGCGCGGAGCGGAGCTGACCGCTTCCGCGAG 638
QY 198 ArgLeuGlyProLeuValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 CGCCTGCGGCGCTGTGTGAGCAGGCGCGCATGTCGCGCGCCACCTGAGTACAGAGGCC 698
QY 218 GlyGlnProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 GGCAGACCCCTGCGCGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 758
QY 238 GluMetGlySerArgThrArgAspArgLeuAspGlnValGlyGlnGlnGlnGlnGlnGln 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 GACATGGCGAGCGCGGAGCCCGGAGCCCGCTGTGATGAGATGCGTACAGCTGAGAGAGTG 818
QY 258 ArgAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 CGCACAAGAGTGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 878
QY 278 ArgLeuGlySerThrPheGluProLeuValGlnAspMetGlnArgGlnTrpAlaGlyLeu 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 879 CTCCTCAAGAGCTGTGCGAGCTGTGTGAGAGACATACGCGGCGAGTGGCGGCGTG 938
QY 298 ValGlnValGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 939 GTGCAAGAGTGCAGTGGCGCTGAGCATTAAGCTCTCCACCTCTGCGCGCATGATTAAT 998

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
Prior APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
Prior APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hieber, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TANK:177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-51

Alignment Scores:
Pred. No.: 2,21e-87 Length: 4267
Score: 1002.00 Matches: 218
Percent Similarity: 58.43% Conservative: 42
Best Local Similarity: 48.99% Mismatches: 46
Query Match: 63.22% Indels: 140
Gaps: 5
US-09-827-854-15 (1-317) x US-08-949-155-51 (1-4267)
QY 1 MetGlyValLeuThrPalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
    ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2448 ATGCGTGTGACAGCTGCTCGT-----GCAGAGTCCGCGAGACAGAGGC 2489
QY 21 GluGlnAlaValGlnThrGluProGluProGluLeuArg-----GlnGlnThrGln 37
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2490 GAGCCGCGG-----CCGCGCGCGAGAGTGCAGCTGTGNGGAGGAGGAGCCAG 2537
QY 38 TTPGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheThrPaspTyrLeuArgTTP 57
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2538 TGGCAGGAGCAGCCAGCCCTGGGAGCAGGCCCTGGGCGCTTGTGGATTAAGTGGCGTG 2597
QY 58 ValGlnThrLeuSerGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2598 GTGCAAGTCCCTGTGTACCAAGTGCAGAGGAGAGCTGTCTACGACCAAGTCAACGAGGA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAGTGCACACCCGAGCTCCCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGCAACCGTGTGTGTGAGCCCTCAGGCTCACCGCGGTTCCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTGTGCCAACCTCTGGGGGTCTGGGTCTGTCTTTCTTTTCTTCTCTCTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAAGCTTTTCTTTTCTTTTCTTTTCAATTTGACTTCATGCTTCTTCC 2896
QY 78 ----- 78

```



```

Db 2837 TTGGGGGAAAAAATTCTTTCTTTCTTTCTTTCAATTGACTTCAATGCTGCTTTCTTTCC 2896
Qy 78 ----- 78
Db 2897 ATCTTAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
Qy 78 ----- 78
Db 2957 AATCTGTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3016
Qy 79 ----- 79
Db 3017 CCCGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3076
Qy 93 LysSerGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 112
Db 3077 CGCGAGAGAGCTGAGAGCGAGCTGAGGCGCCGCTGACCCAGAGAGAGAGAGAGAGAGAGAG 3136
Qy 113 LysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 132
Db 3137 AAGGAGCTGAGAGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3196
Qy 133 LeuValGluThrArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 152
Db 3197 TTGGTCTCTACCGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 3256
Qy 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgLysAspAlaAsp 172
Db 3257 AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3316
Qy 173 LeuGluLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluGluGluGluGluGlu 192
Db 3317 CTGCAAGAGCGCTGCGCGTGTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3376
Qy 193 SerAlaIleArgGluArgLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 212
Db 3377 AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3436
Qy 213 ValGlySerLeuAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 232
Db 3437 CTGAGTACACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3496
Qy 233 ArgAlaArgMetGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 252
Db 3497 CGCGGAGCGCTGAGAGATGTGGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3556
Qy 253 GlnValAlaGluValArgAlaLysLeuGluGluGluGluGluGluGluGluGluGluGlu 272
Db 3557 CAGCTGAGAGAGGTGCGCGCACCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3616
Qy 273 GlnAlaPheGlnAlaArgLeuLysSerTyrPheGluProLeuValGluAspMetGlnArg 292
Db 3617 GAGGATTCACGCGCTCTCAAAAGCGTGTGAGCGCTGTGTGAAGACATACGCGCG 3676
Qy 293 GlnTyrAlaGlyLeuValGluLysValGlnAlaValGly---ThrSerAlaAlaPro 311
Db 3677 CAGTGGCGCGGCTGTGTGAAGATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3736
Qy 312 ValProSerAspAsn 316
Db 3737 GCGCCCATGATAT 3751

```

```

; ADDRESSER: Banner & Wilcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
;
; US-08-726-306A-28
;
; Alignment Scores:
; Pred. No.: 9,84e-87 Length: 660
; Score: 984.00 Matches: 197
; Percent Similarity: 98.99% Conservative: 0
; Best Local Similarity: 98.99% Mismatches: 2
; Query Match: 62.08% Indels: 0
; Gaps: 0
;
; US-09-827-854-15 (1-317) x US-08-726-306A-28 (1-660)
Qy 1 MetLysValLeuTyrPheAlaLeuLeuValTyrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAAGTTCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
Qy 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrPheGln 40
Db 122 GAGCAAGCGGTGAGAGACAGAGCGGAGCGGAGCTGCGCGACAGACGAGTGGCGAGAC 181
Qy 41 GlyGlnArgTyrPheLeuAlaLeuGluGluGluGluGluGluGluGluGluGluGluGlu 60
Db 182 GGCAGCGCTGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGTCTGAGCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluGlu 100
Db 302 CTGATGAGAGAGACATGAAAGAGTGAAGAGCTTCAAAATCGGAATCGAGAGACACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
Qy 121 ArgLeuGlyAlaAspMetGluLysValGlyGlyAlaArgLeuValGlnTyrArgGlyLysVal 140
Db 422 CGGCTGGCGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481

```

QY 141 GlnIlnakettlengugylnglnserThnglunleuAryAlnrgleuAlAserHisleuArg 160
 Db 482 CAGGCCATGCTGGCCGAGACACCGAGAACTCGGGGTGGCCCTGCCCTCCACCTGGCG 541
 QY 161 LysleuArgLysArGLEuLeuArgAspAlaaspserlunLysArGLEuAlaValTyr 180
 Db 542 AAGCTGGCTAAGCGCGCTCCGCCGAGATCCGATGACTCGAAGAGCGCCCGGAGAGTTAC 601
 QY 181 GlnIlnacIyAlaArGgIunGlyAlaGluArgGlyLeuSerSerAlIleArgGluArgLeu 199
 Db 602 CAGCGCCGGGCGCCGCGAGAGGCGCGACGCGGGCTTCAGCGCCATCGCGAGAGCCCTGG 658

RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5525493

```

1  TITLE OF INVENTION:  CLONING METHOD AND KIT
2  NUMBER OF SEQUENCES:  7
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE:  Foley & Lardner
5  STREET:  1800 Diagonal Road, Suite 500
6  City:  Alexandria
7  STATE:  Virginia
8  COUNTRY:  USA
9  ZIP:  22313-0299
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE:  Floppy disk
12 COMPUTER:  IBM PC compatible
13 OPERATING SYSTEM:  PC-DOS/MS-DOS
14 SOFTWARE:  patentin Release #1.0, Version #1.25
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER:  US/07/849,389
17 FILING DATE:  19920519
18 CLASSIFICATION:  435
19 ATTORNEY/AGENT INFORMATION:
20 NAME:  BENT, Stephen A.
21 REGISTRATION NUMBER:  29,768
22 REFERENCE/DOCKET NUMBER:  16787/168/DFBC
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE:  (703)836-9300
25 TELEFAX:  (703)683-4109
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
9
```

Alignment Scores:	
Pred. No.:	2.13e-41
Score:	511.00
Percent Similarity:	98.18%
Best Local Similarity:	97.27%
Query Match:	32.24%
DB:	1
Length:	330
Matches:	107
Conservative:	1
Mismatches:	2
Indels:	0
Gaps:	0

US-09-827-854-15 (1-317) x US-07-849-389-6 (1-330)

Qy 87 LysGluLeuLysAlaTyrLysSerGluLeuGluGluInLeuThrProValAlaGluGlu 106
Db 1 AAGGAGTTGAAGCCTACAAATCGGAACGTGAGCAACAACGTGACCCCGGTGGCGGAGGAG 60

Qy	107	TharnglaaargleuserylgluleuGlnalaaGlnalaaargleuGlyAlaAspMet	120
Db	61	ACGGGGCAGCGCTGCCAAGAGCTCAGGGCGGAGGGCCCCGCTGGCGCGGAATG	120
Qy	127	GluAspValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGln	148
Db	121	GAGGACGTCGGCGCGCTGTGTGTACAGTACCGGGCGAGTGCAGGGCATCTCGGCAG	186
Qy	147	SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgGlyLeuArgGlyArgLeu	166
Db	181	AGCACCGAGGACGTGGCGGTGCGCTGCCCTCCACCTCGCCCAAGCTGCGTAAACGGCTC	240
Qy	167	LeuArgAspAlaAspAspieGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlu	186
Db	241	CTCCGCATGCCATGCACCTCAAAAGCGCTGGCAGTGTACCAAGCGCGGGCCCCGAG	300
Qy	187	GlyAlaGluArgGlyLeuSerAlaIleArg	196
Db	301	GGGCGCGAGCGCGGCTCAGCGGCATCTCGC	330

RESULT 8
US-08-617-256-24
; Sequence 24, Application US/08617256
; Patent No. 6043031

```

1      RESULT 8
2      US-08-617-256-24
3      Sequence 24, Application US/08617256
4      Patent No. 6043031
5      GENERAL INFORMATION:
6      APPLICANT: Kvsler, Hubert
7      TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
8      NUMBER OF SEQUENCES: 33
9      CORRESPONDENCE ADDRESS:
10     ADDRESSEE: LAHYE & COCKFIELD
11     STREET: 60 State Street, suite 510
12     CITY: Boston
13     STATE: Massachusetts
14     COUNTRY: USA
15     ZIP: 02109-1875
16     COMPUTER READABLE FORM:
17     MEDIUM TYPE: floppy disk
18     COMPUTER: IBM PC compatible
19     OPERATING SYSTEM: PC-DOS/MS-DOS
20     SOFTWARE: PatentIn Release #1.0, Version #1.25
21     CURRENT APPLICATION DATA:
22     APPLICATION NUMBER: US/08/617,256
23     FILING DATE: March 18, 1996
24     CLASSIFICATION:
25     PRIOR APPLICATION DATA:
26     APPLICATION NUMBER: 08/406,199
27     FILING DATE: March 17, 1995
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Arnold, Beth A.
30     REGISTRATION NUMBER: 35,430
31     REFERENCE/DOCKET NUMBER: SQT-013CP
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: (617)227-7400
34     TELEFAX: (617)227-5941
35     INFORMATION FOR SEQ. ID NO: 24:
36     SEQUENCE CHARACTERISTICS:
37     LENGTH: 252 base pairs
38     TYPE: nucleic acid
39     STRANDEDNESS: single
40     TOPOLOGY: linear
41     MOLECULE TYPE: CDNA
42     US-08-617-256-24

```

Alignment Scores:	
Pred. No.:	2,63e-28
Length:	252
Score:	374.00
Percent Similarity:	97.62%
Best Local Similarity:	97.62%
Query Match:	23.60%
RB:	3
Gaps:	0
Matches:	82
Conservative:	0
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-854-15 (1-317) x US-08-617-256-24 (1-252)

MOLECULE TYPE: CDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 2,63e-28
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 23.60%

Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1
Gaps: 0

US-09-827-854-15 (1-317) x US-09-431-613-24 (1-252)

QY 109 AAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCGCGCGCTGGCGCGGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGACGTACCGCGCGCGCGCGCATGCTCGCGCAGAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCGCAAGCTGCTGAAGCGGCTCTCCGCG 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGATCCCTGGCAGTGTACAGCGCGGCGCGCGCGCGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCaulliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 2,63e-28
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 23.60%

Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1
Gaps: 0

US-09-827-854-15 (1-317) x US-09-504-245-24 (1-252)

QY 109 AAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCGCGCGCTGGCGCGGACATGGAGAGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGCTGTGTGACGTACCGCGCGCGCGCGCATGCTCGCGCAGAGACACC 120

QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGTGGCTGCTCCACCTGCGCAAGCTGCTGAAGCGGCTCTCCGCG 180

QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGATCCCTGGCAGTGTACAGCGCGGCGCGCGCGCGCGCC 240

QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 12
US-09-287-682-24
Sequence 24, Application US/09287682
Patent No. 6235478
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002E
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 2,63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-287-682-24 (1-252)

QY 109 AlArGrLeuSerLysGluLeuGlnAlaAlaGlnAlaArGrLeuGlnAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArGrLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlnInsThr 148
|||||
DB 62 GTGTGGCC-CGCTGTGGTACGTACCGCGGAGGTGCAGGTCATGCTCGGCCAGAGCACC 120

QY 149 GluGluLeuArGrValArGrLeuAlaSerHisLeuArGlyLeuArGlySarGrLeuLeuArg 168
|||||
DB 121 GAGGAGCTGGGGTGGCCCTCGCTCCACCTCGGCAAGCTGGGTAAAGGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArGrGluGlyAla 188
|||||
DB 181 GATGCCGATACCTGCAGAAAGTCTCGCAGTGTACAGCCGCGGCCCGCGAGGGCGCC 240

QY 189 GluArGlyLeu 192
|||||
DB 241 GACGCGGCGCTC 252

RESULT 13
US-09-287-679-24
Sequence 24, Application US/09287679
Patent No. 6258538
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287, 679
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 2,63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-287-679-24 (1-252)

QY 109 AlArGrLeuSerLysGluLeuGlnAlaAlaGlnAlaArGrLeuGlnAlaAspMetGluAsp 128
|||||
DB 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGCCGCTGGCGCGGACATGAGAGAC 61

QY 129 ValCysGlyArGrLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlnInsThr 148
|||||
DB 62 GTGTGGCC-CGCTGTGGTACGTACCGCGGAGGTGCAGGTCATGCTCGGCCAGAGCACC 120

QY 149 GluGluLeuArGrValArGrLeuAlaSerHisLeuArGlyLeuArGlySarGrLeuLeuArg 168
|||||
DB 121 GAGGAGCTGGGGTGGCCCTCGCTCCACCTCGGCAAGCTGGGTAAAGGGCTCTCCGC 180

QY 169 AspAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArGrGluGlyAla 188
|||||
DB 181 GATGCCGATACCTGCAGAAAGTCTCGCAGTGTACAGCCGCGGCCCGCGAGGGCGCC 240

QY 189 GluArGlyLeu 192
|||||
DB 241 GACGCGGCGCTC 252

RESULT 14
US-09-397-766-24
Sequence 24, Application US/09397766
Patent No. 6268144
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
APPLICANT: Little, Daniel P.
APPLICANT: Braun, Andreas
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```

: APPLICATION NUMBER: US/09/397,766
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/617,256
: FILING DATE: 18-MAR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 24736-20021
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-450-8400
: TELEFAX: 619-587-5360
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-09-397-766-24

Alignment Scores:
Pred. No.: 2.63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-397-766-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGCGGCGGCGGCTGTGGCGGAGCATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGGCTGTGTGAGTACCGCGGCGAGGTGCAAGCCATCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTCCGCTCCCTCCACCTGCGCAAGCTGCTGAAGCGGCTCCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGGCGGCGCGGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252

RESULT 15
US-09-287-681-24
: Sequence 24, Application US/09287681
: Patent No. 6277573
: GENERAL INFORMATION:
: APPLICANT: K sier, Hubert
: TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Heller Ehrman White & McCauliffe
: STREET: 4250 Executive Square, 7th Floor
: CITY: La Jolla
: STATE: CA
: COUNTRY: USA
: ZIP: 92037-9103
```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/287,681
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/617,256
: FILING DATE: 18-MAR-1996
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/406,199
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 24736-2002F
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-450-8400
: TELEFAX: 619-587-5360
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)227-5941
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 252 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-09-287-681-24

Alignment Scores:
Pred. No.: 2.63e-28 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 23.60% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15 (1-317) x US-09-287-681-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGCGGCGGCGGCTGTGGCGGAGCATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGCGC-CGGCTGTGTGAGTACCGCGGCGAGGTGCAAGCCATCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTGCGGGTCCGCTCCCTCCACCTGCGCAAGCTGCTGAAGCGGCTCCTCCGC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGluAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTGTACCAAGGCGGCGCGGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCGCTC 252
```

Search completed: March 14, 2003, 20:18:27
Job time : 39.7343 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 65.8753 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15
Perfect score: 1585
Sequence: 1 MKVLMALVTLFLAGCOAKV.....VEKVOAAVGTSAFVPSDNH 317

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 segs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xip
-O=/cgn2_1/USPTO.spool/US09827854/runat_11032003_101612_27578/app_query.fasta.1.3576
-DB=Published.Applications_NA -OPMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOEXT=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09827854.ecgn_1_1.298.4runat_11032003_101612.27578
-NCPU=6 -ICPU=3 -NO_XLPTX -NO_MAP -LARGEOUTER -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications.NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PTC_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	1585	100.0	1156 9	US-09-870-759-129	Sequence 129, App
2	1585	100.0	1156 9	US-09-802-640-17	Sequence 17, Appl
3	1585	100.0	1156 10	US-09-827-854-8	Sequence 8, Appl
4	1585	100.0	1291 12	US-10-044-090-454	Sequence 454, App

5	1581	99.7	1156 10	US-09-827-854-12	Sequence 12, Appl
6	1577	99.5	1156 10	US-09-827-854-9	Sequence 9, Appl
7	1577	99.5	1156 10	US-09-827-854-11	Sequence 11, Appl
8	1576	99.4	1157 10	US-09-954-456-760	Sequence 760, App
9	1576	99.4	1157 10	US-09-880-107-2244	Sequence 2244, App
10	1573	99.2	1156 10	US-09-827-854-7	Sequence 7, Appl
11	1570	99.1	1156 10	US-09-827-854-10	Sequence 10, Appl
12	1369.5	86.4	41907 10	US-09-967-013-5	Sequence 5, Appl
13	867.5	54.7	786 10	US-09-925-302-133	Sequence 133, App
14	625	39.4	478 10	US-09-964-824-374	Sequence 374, App
15	625	39.4	478 10	US-09-880-107-2491	Sequence 2491, App
16	463	29.2	356 10	US-09-960-352-5420	Sequence 5420, App
17	458.5	28.9	446 10	US-09-960-352-4726	Sequence 4726, App
18	439.5	27.7	442 10	US-09-960-352-9395	Sequence 9395, App
19	435	27.4	414 10	US-09-960-352-4237	Sequence 4237, App
20	428.5	27.0	416 10	US-09-960-352-8720	Sequence 8720, App
21	422.5	26.7	423 10	US-09-960-352-14047	Sequence 14047, A
22	405.5	25.6	409 10	US-09-960-352-5148	Sequence 5148, App
23	401.5	25.3	416 10	US-09-960-352-9797	Sequence 9797, App
24	401	25.3	253 10	US-09-179-5368-130	Sequence 130, App
25	387.5	24.4	425 10	US-09-960-352-3497	Sequence 3497, App
26	380.5	24.0	390 10	US-09-960-352-1311	Sequence 1311, App
27	379.5	23.9	391 10	US-09-960-352-1278	Sequence 1278, App
28	379.5	23.9	392 10	US-09-960-352-10599	Sequence 10599, A
29	379.5	23.9	401 10	US-09-960-352-4187	Sequence 4187, App
30	375.5	23.7	413 10	US-09-960-352-8042	Sequence 8042, App
31	374	23.6	252 10	US-09-796-116-24	Sequence 24, Appl
32	374	23.6	252 10	US-09-879-341-24	Sequence 24, Appl
33	369.5	23.3	404 10	US-09-960-352-2325	Sequence 2325, App
34	369.5	23.3	405 10	US-09-960-352-14063	Sequence 14063, A
35	351.5	22.2	377 10	US-09-960-352-176	Sequence 176, App
36	350.5	22.1	353 10	US-09-960-352-4914	Sequence 4914, App
37	350.5	22.1	378 10	US-09-960-352-10040	Sequence 10040, A
38	350.5	22.1	388 10	US-09-960-352-10361	Sequence 10361, A
39	348.5	22.0	388 10	US-09-960-352-11986	Sequence 11986, A
40	345.5	21.8	396 10	US-09-960-352-3457	Sequence 3457, App
41	344.5	21.7	377 10	US-09-960-352-10540	Sequence 10540, A
42	338.5	21.4	432 10	US-09-960-352-8934	Sequence 8934, App
43	338	21.3	332 10	US-09-960-352-6552	Sequence 6552, App
44	337.5	21.3	376 10	US-09-960-352-5408	Sequence 5408, App
45	335.5	21.2	373 10	US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-870-759-129
Sequence 129, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, DAVID S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT FILING DATE: 2002-01-14
PRIORITY APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 129
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION:
US-09-870-759-129
Alignment Scores:
Pred. No.: 4.99e-136
Score: 1585.00
Percent Similarity: 100.00%
Length: 1156
Matches: 317
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-827-854-15 (1-317) x US-09-870-759-129 (1-1156)

QY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 61 ATGAAGTTCTGTGGCTGGTGTGCTGTCACATCTCTGACAGATCCAGGCAAGTGT 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPglInSer 40
 |||||||
 Db 121 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGCGCAGCAGACGAGGAGGAGC 180
 QY 41 GlyLysArgThrPgluLeuAlaLeuGlyArgPheThrPaspTyrLeuArgTyrPValGlnThr 60
 |||||||
 Db 181 GGCAGCGCTGGGAACATGGACATGGCTGCTTTGGGATTACCTGGCTGGTGGTCAACA 240
 QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 Db 241 CTGTCTGAGCAGGTGCGAGGAGGAGCTGCTCAGCTCCAGGTCAACCCAGAACTGAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 Db 301 CTGATGACACAGACCAATGAAGAGCTTAAGCCCTACAAATCGAGACTGGAGCAACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||||
 Db 361 ACCCGCTGGCGAGAGGAGCGGGCGGCTGCCAAGAGCTGACGCGGCGGCGAGGGC 420
 QY 121 ArgLeuGlnLysAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 421 CGGCTGGCGCGGACATGAGGAGAGCTGTGCGCGCTGTGAGTACCCGCGGAGAGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 481 CAGGCATGCTGGCGCAGACACGAGAGACTCGGGTGGCGCTCGCTCCACACTGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 541 AACCTGCTGAAGCGGCTCCCGCATGCGCATGACCTGCAGAAAGCCCTGGCAGATGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlnGluArgLysGlyLeuSerLysAlaThrGluValArgLeuGly 200
 |||||||
 Db 601 CAGGCGGGGCGCGGAGGCGGCGGAGCGGCGCTCAGCGCATCCGCGAGCGCTGGGG 660
 QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
 |||||||
 Db 661 CCCTGTGTGAACAGGCGCGCTGCGGGCGGCGCACTGTGGCTCTCGGCGGCGAGCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaTyrPglLysGluArgLeuArgAlaArgMetGluMetGly 240
 |||||||
 Db 721 CTACAGAGAGGCGGCGGCTGCGGCGGAGCGGCTGCGCGCGGATGAGAGATGGGG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValAlaGlnAlaLys 260
 |||||||
 Db 781 AGCGGAGCCGCGCAGCGCTGAGGAGGTGAAGAGCAGAGTGGCGAGGGCGCGGCAAG 840
 QY 261 LeuGlnGluGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
 |||||||
 Db 841 CTGGAGAGAGGCGCAGCAGATACCGCTGCGAGGCGGAGGCGCTTCCAGGCGCGCTCAAG 900
 QY 281 SerTyrPheGluProLeuValGlnAspMetGlnArgGlnThrPAlaGlnLysValGlnLys 300
 |||||||
 Db 901 AGCTGTGTGAGAGCCCTGTGGGAAAGATGACAGCGCAGTGGCGGCGGTGGGAAAG 960
 QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
 |||||||
 Db 961 GTGCAGAGCTGCGTGGGCAACAGCGCGCGCTGTGCGCAGCAGCAATCAAC 1011
 RESULT 2
 US-09-802-640-17
 ; Sequence 17, Application US/09802640

; Publication No. US20030036057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Braun, Andreas
 ; APPLICANT: Bonsal Aruna
 ; APPLICANT: Kleyn Patrick
 ; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: 24736-2048
 ; CURRENT APPLICATION NUMBER: US/09/802,640
 ; NUMBER OF SEQ ID NOS: 122
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 1156
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1014)
 ; OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E
 ; OTHER INFORMATION: (APOE)
 US-09-802-640-17

Alignment Scores:
 Pred. No.: 4,99e-136 Length: 1156
 Score: 1585.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-827-854-15 (1-317) x US-09-802-640-17 (1-1156)

QY 1 MetLysValLeuThrPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 Db 61 ATGAAGTTCTGTGGCTGGTGTGCTGTCACATCTCTGACAGATCCAGGCAAGTGT 120
 QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPglInSer 40
 |||||||
 Db 121 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGCGCAGCAGACGAGGAGGAGC 180
 QY 41 GlyLysArgThrPgluLeuAlaLeuGlyArgPheThrPaspTyrLeuArgTyrPValGlnThr 60
 |||||||
 Db 181 GGCAGCGCTGGGAACATGGAGAGCTGTGCGCGCTGTGAGTACCCGCGGAGAGTG 240
 QY 61 LeuSerGluGlnValGlnGlnArgLysGlnValThrGlnLysLeuArgAla 80
 |||||||
 Db 241 CTGTCTGAGCAGGTGCGAGGAGGAGCTGCTCAGCTCCAGGTCAACCCAGAACTGAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||||
 Db 301 CTGATGACACAGACCAATGAAGAGCTTAAGAGCCCTACAAATCGAGACTGGAGCAACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerSerGlnValThrGlnLysLeuArgAla 120
 |||||||
 Db 361 ACCCGCTGGCGAGAGGAGCGGCGGAGCGGCTGTCCAAAGAGCTGAGCGGCGGCGGCGC 420
 QY 121 ArgLeuGlnLysAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 Db 421 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCGCTGTGTCAGTACCGCGGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 Db 481 CAGGCATGCTGGCGCAGACACCGGAGAGCTGCGGGTGGCGCTGCTCCACAGCGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 Db 541 AACCTGCTGAAGCGGCTCCCGCATGCGCATGACCTGCAGAAAGCCCTGGCAGATGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlnGluArgLysGlyLeuSerLysAlaThrGluValArgLeuGly 200
 |||||||
 Db 601 CAGGCGGGGCGCGGAGGCGGCGGAGCGGCGCTGACGCCATCCGCGAGCGGCTGGGG 660

Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAACAGAGCGCGCTGGCGGCCCACTGTGGGCTCCCTCGCGCGCCAGCCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGGAGCGGCGCCAGGCTGGGGCGAGCGGCTGGCGCGCGATGAGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 761 AGCCGACCCCGAGCGCGCTGGAGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGTTGAGAGCCCTGGTGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGAGGCTGCGGTGGGACAGCGCCGCTGTGGCCAGCAGCAATCAC 1011

RESULT 3
US-09-827-854-8
: Sequence 8, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Zannis, Vassilis
: APPLICANT: Kyriacos, Kyriacos E.
: TITLE OF INVENTION: Compounds and methods for lowering
: TITLE OF INVENTION: cholesterol levels without inducing hypertiglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827,854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679,088
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 09/544,386
: PRIOR FILING DATE: 2000-04-06
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 1156
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-827-854-8

Alignment Scores:
Pred. NO.: 4,99e-136 Length: 1156
Score: 1585.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15 (1-317) x US-09-827-854-8 (1-1156)

Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCGCTGTGGCTGTGCATTCCTGGAGAGATGCCAGGCAAGGTG 120
Qy 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCCGAGCGCGCCAGCAGACCGAAGTGGCAGAGC 180
Qy 41 GlyLysArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLysArgTrpValGlnThr 60
Db 181 GAGCAGCGCTGGAGACAGAGCGCGCTGGGTGGTGGATTCCTGGCGGTGGGTGGAGCA 240
Qy 61 LeuSerGluGlnValGlnGluGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 241 CTGCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGCAAGTGAAGGCG 300

Qy 81 LeuMetAspGluTrpMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGAGACAGACCATTAAGAGAGTGAAGGCTTACAAATGGAACTGGAGAACTG 360
Qy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGGAGAGACAGCGGCGAGCGCTGTCTCAAGAGAGCTGACAGCGCGAGCG 420
Qy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACATGAGAGAGAGTGTGCGCGCGCTGTGGCTAGTACCGCGGAGAGTG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGluGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGCCATGCTGGCGCAGAGCAGCAGAGAGTGGCGGCTGCGCTCCACCTCGCG 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
Db 541 AAGCTGCGTAAAGCGGCTCTCTCGGAGTCCGATGACCTGCAGAAAGCGCTGTGCAGTAC 600
Qy 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGCGCGCGAGAGCGCGAGCGCGCTCAGCGCCATCCGAGCGCGCTGGGG 660
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAACAGAGCGCGCTGGCGGCCCACTGTGGGCTCCCTGGCGCGCCAGCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCGCCAGGCTGGGGGAGAGCGGTGGCGCGCGGAGAGAGAGATGGGC 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 781 AGCCGAGCGCGCGAGCGCGCTGGAGAGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAG 900
Qy 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGTTGAGAGCCCTGGTGAAGACATGCAGCGCCAGTGGCGCGGCTGTGGAGAG 960
Qy 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 961 GTGAGGCTGCGGTGGGACAGCGCCGCTGTGGCCAGCAGCAATCAC 1011

RESULT 4
US-10-044-090-454
: Sequence 454, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-0028 US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO 454
: LENGTH: 1291
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. US20020137081A1 2514988CB1
: NAME/KEY: unsure
: LOCATION: 46
: OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-454

US-09-827-854-11

Alignment Scores:

Pred. No.:	2,68e-135	Length:	1156
Score:	1577.00	Matches:	316
Percent Similarity:	99.68%	Conservative:	0
Best Local Similarity:	99.68%	Mismatches:	1
Query Match:	99.50%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15 (1-317) x US-09-827-854-11 (1-1156)

```
QY 1 MettysValleuThrpAlaalaLeuValThPhleuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGGCTGGTGGTCTGTGCACATTCTGTGCAGAGATCCAGGCCAAAGTG 120
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGAGACGAGAGTGAGAGC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGGCAGGCGTGGAGACGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTGTAGCAGGTGGAGAGAGCTGTGCAGCTCCAGGTCCAGGAGAGAGAGAGAGAG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGGACAGACCATGAAGAGATTGAAGCCATCAAAATGGAACTGGAGAACTG 360
QY 101 ThProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGCGGGGACGGCTGTCCAAAGAGCTGCAGGGGCGAGAGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
Db 421 CGGCTGGGCGCGACATGAGAGAGCTGTGGCGGCTGTGTGACAGTACCCCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCAGACACCGAGAGCTGGGGTGGCTCCCTCCACCTGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AACCTGTGAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCCCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
Db 601 CAGCGCGGGGCGCGAGGCGCGAGCGCGGCTCAGCCCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
Db 661 CCCCTGTGGAACAGGGCCCGGTGGCGGCCGACATGTGGCTCCCTGGCGCGCCAGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly 240
Db 721 CTACAGAGAGCGGCCAGGCTGTGGGGAGCGGCTGCGCGCGGAGATGAGAGATGGCG 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGGGAGCGCGCAGCGCTGTGAGCGAGGTGAAAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnAlaLeuArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 841 CTGAGAGAGAGCGCCAGCGATACGCTGTGAGCGAGCGAGCGCTCCAGGGCCCGCTCAAG 900
QY 281 SerTrpPheGluProLeuValGluAspMetGlnArgGlnTrpAlaGlyLeuValGlnLys 300
Db 901 AGCTGTGTGAGCGCCCTGTGTGGAAGACATGACAGCGCAGTGGCGCGGTGGTGGAGAG 960
QY 301 ValGlnAlaAlaValGlyThrSerAlaAlaProValProSerAspAsnHis 317
```

```
Db 961 GTGAGGCTGCGGTGGGACCGCCGCTGTGTGCCAGCAATTCAC 1011
RESULT 8
US-09-954-456-760
; Sequence 760, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 760
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-954-456-760
Alignment Scores:
Pred. No.: 3.3e-135 Length: 1157
Score: 1576.00 Matches: 315
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 2
Query Match: 99.43% Indels: 0
DB: 10 Gaps: 0
US-09-827-854-15 (1-317) x US-09-954-456-760 (1-1157)
QY 1 MettysValleuThrpAlaalaLeuValThPhleuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCTGTGGCTGGTGGTCTGTGCACATTCTGTGCAGAGATCCAGGCCAAAGTG 121
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
Db 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGAGACGAGAGTGAGAGC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGGCAGGCGTGGAGACGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGCTGTAGCAGGTGGAGAGAGCTGTGCAGCTCCAAAGTCCCAAGAACTGAGAGCGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 302 CTGATGGACAGACCATGAAGAGATTGAAGCCATCAAAATGGAACTGGAGAACTG 361
QY 101 ThProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
```

```

|||||
Db 362 ACCCGGTAGCGAGGAGACCGCGGCGCTGCTCCAGAGAGCTGCACAGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValIcysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGCGGAGCATGAGACCTGTGGCGCCCTGGTGCTACATCCGGCGGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTGGCCACAGACACCGAGAGCTGGGGTGGCTTCCCTCCACCTGGCC 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCTAAGCGGCTCTCCGCGATCCGATCGCTGACAGAAACGGCTGGCAGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 602 CAGGCGGGGGCGCGGAGGGGGCGCGAGCGCGCTCAGCGCCATCCGGAGCGCTGGGG 661
Qy 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGGAAACAGGGCGCGGTGGGGCGCGCACTGTGGCTCCCTGGCGCGCAGCGC 721
Qy 221 LeuGlnGluArgAlaGlnAlaTyrGlyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 722 CTACAGAGCGGGCGCGCGCTGGGGCGAGCGGCTGGCGCGCGGATGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGACCGCGGACCGCTGTGAGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCGCCAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCAGCGCCACAGATACGCTGCAGCGCGGAGGCTTCCAGCGCCCTCAAG 901
Qy 281 SerTyrPheGluProLeuValGlnAspMetGlnArgGlnTyrPalaGlyLeuValGlnLys 300
Db 902 AGCTGCTTCGAGCGCCCTGTGTAAGACATGCAGCGCAGTGGCGGCTGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCGAGCTGCGGTGGGACACAGCGCCCGCTGTGTGCCAGCGACAAATCAC 1012

RESULT 9
US-09-880-107-2244
: Sequence 2244, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2244
: LENGTH: 1157
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

```

```

Percent Similarity: 99.37%      Conservative: 0
Best Local Similarity: 99.37%      Mismatches: 2
Query Match: 99.43%      Indels: 0
DB: 10      Gaps: 0

US-09-827-854-15 (1-317) x US-09-880-107-2244 (1-1157)
Qy 1 MetLysValLeuTyrPalaAlaLeuLeuValTyrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTCCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 121
Qy 21 GlnGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnTyrPalaLys 40
Db 122 GAGCAAGGGGTGAGACACAGAGCGGAGCCGAGCTGGCCACAGACAGACGAGTGGAGAGC 181
Qy 41 GlyGlnArgTyrPalaAlaLeuGlnLysArgPheTyrPalaLysArgTyrPalaLys 60
Db 182 GCGCAGCGCTGGGAACCTGGCACTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGACA 241
Qy 61 LeuSerGlnGlnValAlaGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 242 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCCAAGTACCCCAAGAACTGAGGGCG 301
Qy 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 302 CTGATGAGCAGAGACCATGAGAGATTGAAGGCTTCAAAATGGAACTGGAGAACACTG 361
Qy 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGAGAGAACAGCGGGCGAGCGGTCTCAAGAGACTGCACAGCGCGCAGGCC 421
Qy 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGGGCGCGAGCATGAGAGACTGTGGCGCCCTGTGTCAATACCGCGCGCAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGlnLeuLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTGGCGCAGAGACACAGAGAGCTGGGGTGGCGCTGCCCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCTAAGCGGCTCTCCGCGATCCCGATGCCATGCACAGAAACGGCTGGCAGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 602 CAGCGCGGGCGCGCGAGGGCGCGCAGCGCGGCTTCAAGCGCATCCGAGCGGCTGGGG 661
Qy 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 662 CCCCTGGTGGAAACAGGGCGCGGTGGGGCGCGCACTGTGGGCTCCCTGGCGCGCAGCGC 721
Qy 221 LeuGlnGluArgAlaGlnAlaTyrGlyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 722 CTACAGAGCGGGCGCGCGCTGTGGGCGAGCGGCTGGCGCGGATGAGAGATGGGC 781
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 782 AGTCGACCGCGGACCGCTGTGTAAGACATGCAGCGCAGTGGCGGCTGGTGGAGAG 841
Qy 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAlaArgLeuLys 280
Db 842 CTGGAGAGAGCAGCGCCACAGATACGCTGCAGCGCGGAGGCTTCCAGCGCGCTCAAG 901
Qy 281 SerTyrPheGluProLeuValGlnAspMetGlnArgGlnTyrPalaGlyLeuValGlnLys 300
Db 902 AGCTGCTTCGAGCGCCCTGTGTAAGACATGCAGCGCAGTGGCGGCTGGTGGAGAG 961
Qy 301 ValGlnAlaAlaValAlaGlyThrSerAlaAlaProValProSerAspAsnHis 317
Db 962 GTGCGAGCTGCGGTGGGACACAGCGCCCGCTGTGTGCCAGCGACAAATCAC 1012

RESULT 10
US-09-827-854-7

```

Alignment Scores: 3.3e-135 Length: 1157
Pred. No.: 1576.00 Matches: 315

Oy	158	sleuArTgYsleuAaVgYsArSleuLeuAaTgspAaPaPaPleuAInLysArSleuA	178
Db	21336	CCTGGCGAAGCTGtGTAAAGCGGCTCTCCCGCATGCTGCACAGAAAGTGGCTGGC	21395S
Oy	178	aValTgYrGlnAlaGlaYlaAaTgUglYAlaGluAaTgYleuSerAlaIleArGluAa	198
Db	21396	AGTGTACCAAGCGCGGGGCCCGGAGAGCGGCCAGAGCGGGCCCTACAGGCCATCGGAGAGC	21455S
Oy	198	gLeuGlyProLeuValGluGlnGlyArGValAaTgAlaAaIaThrValGlySerLeuAlaG	218
Db	21456	CCTGGGGCCCCCTGGGTGGAAACAGAGCGGGGTGGGGGCCCGCATGTGGGTCTCCCTGGCGG	21515S
Oy	218	YgInProLeuGlnGluAaTgAlaGlnAlaAaTgPcYgLuAaTgLeuAaTgArGAlaAaTgMetGluG	238
Db	21516	CCAGCCGCTACAGAGAGCGGGGCCCAAGGCTTGGGGGACAGCGGCTGGCCCGCGCATGTGAAGA	21575S
Oy	238	uMetGlySerAaTgThrAaTgAsPaTgLeuAspGluValIysGluGlnValAlaGluAa	258
Db	21576	GATGGGCGAGCCGAGACCCCGGACCGCTGTGAGAGGTGAAGGAGCAGGTGGCGGAGGTGGC	21635S
Oy	258	gAlaIaTgLeuGluGluGlnAlaGlnGlnIleAaTgLeuGlnAlaGluAlaAaPheGlnAlaAa	278
Db	21636	CGCCAGAGCTGAGAGAGAGCGCCCAACAGATPACGCTGCAGAGCGGAGGCGCTTCCAGAGGCCG	21695S
Oy	278	gLeuYsSerTgPPaGluProLeuValGluAspMetGlnArGInTgPAlaGlyLeuVa	298
Db	21696	CCTCAAGAGCTGTGTCGAGCCCTGTGTGAACATGACATGCAGCGCCCAATGGCGCCGGGCTGT	21755S
Oy	298	IGluIySValGlnAlaAaIaValAlGlyThSerAlaAaProValProSerAspAsnHis	317
Db	21756	GGAGAAAGGTGCGAGGCTGGCCGTGGGACCAAGCGCCGCCCTGTGGCCACGACAAATCAC	21813
RESULT 13			
US-09-925-302-133			
; Sequence 133: Application US/09925302			
; Patent No. US20020044941A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies			
; FILE REFERENCE: PA104			
; CURRENT APPLICATION NUMBER: US/09/925,302			
; CURRENT FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: PCT/US00/05918			
; PRIOR FILING DATE: 2000-03-08			
; PRIOR APPLICATION NUMBER: 60/124,270			
; PRIOR FILING DATE: 1999-03-12			
; NUMBER OF SEQ ID NOS: 896			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 133			
; LENGTH: 786			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-925-302-133			

Alignment Scores:	
Pred. No.:	7,81e-71
Score:	867.50
Percent Similarity:	98.39%
Best local Similarity:	98.39%
Query Match:	54.73%
DB:	10
	Gaps:
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:

Qy	132	ArgLeuValGlnTyrArgGlyGlnValGlnAlaMetLeuGlnGlnSerThrGlnGlnLeu	151
Db	6	CGCCGGCGACGTACCGGGGAGAGTCGASCCATTGCTGGCCAGACACCGAGACTG	65
Qy	152	ArgValAlaArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp	171
Db	66	CGGGTGGCCCTGGCTCCACACTTCGGGAA-CTGGGTAA-CGGCTCCCTCCGAGATGGCGGAT	123

OY	172	AspleuGlnTysArgLeuAlaIleValTyrGlnAlaGlyAlaArgIuGluAlaGluArgGly	191
Db	124	GACCTCGAAGACCCCTGCGAGTGTACACAGCCCGGGCCGCCAGAGGCCCGCAGCGCGGC	183
OY	192	LeuSerAlaIleIeaArgGluIuArgLeuGlyProLeuValGluGlnGlyArgValaArgAlaIa	211
Db	184	CTCAGCGCCATCCCGCGAGCCCTGTGGGCCCTGTGTGGAMACAGGGCCCGCTGGCGGCCCC	243
OY	212	ThrValGlySerLeuAlaGlyGlnProLeuGlnGluArgGlnAlaIaTyrPglYglIuArg	231
Db	244	ACTGTGGGCTCTCCCTGGCCGGCGACCCCTACAGAGCGGGCCCAAGGCTTGGCGGAGCGG	303
OY	232	LeuArgAlaIaArgMetGluGluMetGlySerArgThrArgAspArgLeuAspGluValTys	251
Db	304	CTGGCGCGCGGAGTGGAGAGATGGGACGCGGACCCGCGACCGCCTGGACGAGGTGAG	363
OY	252	GluGlnValAlaGluValaArgAlaIleLysLeuGlnGluGlnAlaGlnGlnIleArgLeuGln	271
Db	364	GAGCAGGTGGCGGAGGTGGCGCGCCAACTGTGAGAGCAGGCCCCAGCAGATACCCCTGCAG	423
OY	272	AlaGluAlaIlePheGlnAlaArgLeuLysSerTyrPheGluProLeuValaGluAspMetGln	291
Db	424	GCCGAGGCGCTTCCAGGCGCCGCGCTCAAAAGCTGTTCCAGGCCCCCTGTGTGAAGACATGCAG	483
OY	292	ArgGlnTrrPalaGlyLeuValaGluLysValaGlnAlaIleAlaValaGlyThrSerAlaIaPro	311
Db	484	CGCCAGTGGCGCGCGGTGTGTGAAGAAAGTGCAGAGCTGCCCTGTGGGCACACAGCCGCCCT	543
OY	312	ValProSerAspAsnHis	317
Db	544	GTGGCCACGACATTCAC	561

```

RESULT 14
US-09-964-824A-374/C
: Sequence 374, Application US/09964824A
: Patent No. US20020102531A1
: GENERAL INFORMATION:
: APPLICANT: Horrigan, Stephen
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-73
: CURRENT APPLICATION NUMBER: US/09/964, 824A
: CURRENT FILING DATE: 2001-09-27
: PRIOR APPLICATION NUMBER: US/60/236,033
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US/60/236,032
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US/60/236,028
: PRIOR FILING DATE: 2000-09-28
: NUMBER OF SEQ ID NOS: 583
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 374
: LENGTH: 478
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(478)
: OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

```

Alignment Scores:	
Prod. No.:	5.45e-49
Score:	625.00
Percent Similarity:	97.92%
Best Local Similarity:	96.62%
Query Match:	39,438
DB:	10
Length:	478
Matches:	126
Conservative:	1
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-854-15 (1-317) x US-09-964-824A-374 (1-478)

Qy 1 MettysValLeuTrpAlaIaLeuValIhrPhenAlaGlyCysGlnAlaVal 20
|||||

```

Db 396 ATGAGGTTCTGTGGCTGCTGTGCTACATTCTCGGACGATGCCAGGCAAGGTG 337
Qy 21 GtGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGACCGAGCCNCAGCTGCGCCAGACAGACGAGTGGCAGAGC 277
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGluYarPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACGCGCTGGAACTGGCACTGGGTGCTTTGGATTACCTGCTGGGTGGAGACA 217
Qy 61 LeuSerGluGlnValGlnGluGluLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGCTGTAGCAGGTGCAGAGAGAGAGCTGCTCCAGCTCCAGCTCAGCACTGAGAGGC 157
Qy 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLe 100
Db 156 GCTGATGACGACACCATGAAGAGGTTGAAGCCCTACAAATCGGAACGAGGAACAACCT 97
Qy 100 uThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGCGGAGAGAGACGGCGGCTGTCCAGAGAGCTGCAGGCGGCGGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGCGCGGACATGAGAGACGTG 9

```

RESULT 15

```

US-09-880-107-2491/c
: Sequence 2491, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scheff, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-NO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2491
: LENGTH: 478
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 N33009
: NAME/KEY: unsure
: LOCATION: (1)..(478)
: OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2491

```

Alignment Scores:

```

Pred. No.: 5.45e-49
Score: 625.00
Percent Similarity: 97.69%
Best Local Similarity: 96.92%
Query Match: 39.43%
DB: 10
Gaps: 0

```

US-09-827-854-15 (1-317) x US-09-880-107-2491 (1-478)

```

Qy 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 396 ATGAGGTTCTGTGGCTGCTGTGCTACATTCTCGGACGATGCCAGGCAAGGTG 337
Qy 21 GtGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 336 GAGCAAGCGGTGGAGACAGACCGAGCCNCAGCTGCGCCAGACAGACGAGTGGCAGAGC 277

```

```

Qy 41 GlyGlnArgTrpGluLeuAlaLeuGluYarPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 276 GGCACGCGCTGGAACTGGCACTGGGTGCTTTGGATTACCTGCTGGGTGGAGACA 217
Qy 61 LeuSerGluGlnValGlnGluGluLeu-SerSerGlnValThrGlnGluLeuArgAl 80
Db 216 CTGCTGTAGCAGGTGCAGAGAGAGAGCTGCTCCAGCTCCAGCTCAGCACTGAGAGGC 157
Qy 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLe 100
Db 156 GCTGATGACGACACCATGAAGAGGTTGAAGCCCTACAAATCGGAACGAGGAACAACCT 97
Qy 100 uThrProValAlaGluGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAl 120
Db 96 GACCCCGGTGGCGGAGAGAGACGGCGGCTGTCCAGAGAGCTGCAGGCGGCGGAGN 37
Qy 120 aArgLeuGlyAlaAspMetGluAspVal 129
Db 36 CCGGCTGGCGCGGACATGAGAGACGTG 9

```

Search completed: March 15, 2003, 03:03:43
Job time : 80.8753 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1289.69 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15

Perfect score: 1385
Sequence: 1 MKVLMALVTLFLAGCAKY.....VEKVOAAVGTSAAPVPSDNH 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_plus_p2n.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US09827854/funat_11032003_101610_27495/app_query.fasta.1.3576  
-DB=EST -QFMT=fastest -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09827854.gcgn.1.1.8826.grunat.11032003_101610_27495 -NCPU=6 -ICPU=3  
-NO_XLPT -NO_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*  
1: em_estba:*  
2: em_esthum:*  
3: em_estlin:*  
4: em_estlmu:*  
5: em_estlov:*  
6: em_estlpl:*  
7: em_estlro:*  
8: em_hlc:*  
9: gb_est1:*  
10: gb_est2:*  
11: gb_hlc:*  
12: gb_est3:*  
13: gb_est4:*  
14: gb_est5:*  
15: em_estfun:*  
16: em_estom:*  
17: gb_gss:*  
18: em_gss_hum:*  
19: em_gss_inv:*  
20: em_gss_pln:*  
21: em_gss_vit:*  
22: em_gss_fun:*  
23: em_gss_mam:*  
24: em_gss_mus:*  
25: em_gss_other:*  
26: em_gss_pro:*  
27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query		ALIGNMENTS	
No.	Score	Match	Length	DB	ID
1	1290.5	81.4	1027	13	B1670367
2	1254	79.1	842	12	BG763371
3	1248	78.7	933	13	B1668318
4	1220	77.0	938	12	BG761746
5	1217.5	76.8	942	13	B1600906
6	1203	75.9	800	13	BM042094
7	1202.5	75.9	922	12	BG472299
8	1199.5	75.7	922	13	B1597743
9	1190.5	75.1	817	12	BG774871
10	1189	75.0	845	12	B1600563
11	1175.5	74.2	845	12	BG829472
12	1152	72.7	790	12	BG707147
13	1149	72.5	919	13	B1551475
14	1139	71.9	757	13	BM042228
15	1129.5	71.3	907	12	BG706129
16	1128	71.2	706	14	BM728696
17	1128	71.2	741	12	BG762924
18	1126	71.0	1110	11	AK010261
19	1124	70.9	804	12	BG702752
20	1123	70.9	834	13	BM042676
21	1120.5	70.7	808	13	B1668329
22	1117	70.5	855	13	B161362
23	1114	70.3	803	13	B1670350
24	1114	70.1	782	12	BG716776
25	1111	70.1	812	13	B1601551
26	1111	70.1	965	14	BQ677266
27	1108	69.9	757	13	B1603658
28	1105.5	69.7	797	12	BG715366
29	1105	69.7	846	13	B1159757
30	1104	69.7	796	13	BM042153
31	1099	69.3	757	12	BG707750
32	1099	69.3	794	13	B1601279
33	1099	69.3	954	12	BF967543
34	1097	69.2	790	13	B1551066
35	1096	69.1	790	13	B1551811
36	1092.5	68.9	798	12	BG708414
37	1091.5	68.9	891	13	B1549222
38	1085.5	68.5	802	13	B1458355
39	1084	68.4	1100	14	BM914382
40	1083.5	68.4	914	13	B1603523
41	1083	68.3	748	13	B1553085
42	1082.5	68.3	930	12	BF967857
43	1078.5	68.0	812	12	BG769968
44	1078	68.0	688	13	B1544886
45	1077.5	68.0	803	12	BG709360

ALIGNMENTS

```
RESULT 1  
B1670367  
LOCUS      B1670367      1027 bp      mRNA  
DEFINITION 603292738P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5312024 5',  
LOCUS      mRNA sequence.  
ACCESSION  B1670367  
VERSION    B1670367.1 GI:15584600  
KEYWORDS   EST.  
SOURCE      human.  
ORGANISM   Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  1 (bases 1 to 1027)  
NIH-MGC http://mgc.nci.nih.gov/.  
TITL        National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshitsuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11790 row: 1 column: 09
 High quality sequence stop: 845.
 Location/Qualifiers

FEATURES

source

```
1..1027
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5312024"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
```

BASE COUNT 194 a 308 c 400 g 125 t
 ORIGIN

Alignment Scores:

Pred. No.: 1,41e-125 Length: 1027
 Score: 1290.50 Matches: 292
 Percent Similarity: 90.85% Conservative: 6
 Best Local Similarity: 89.02% Mismatches: 19
 Query Match: 81.42% Indels: 12
 DB: 13 Gaps: 2

US-09-827-854-15 (1-317) x B1670367 (1-1027)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 DB 26 ATGAAGGTTCTGTGGCTGTGTGTCACATTCCTGCAGAGATGCCAGCCAGATG 85
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluProGln 40
 |||||||
 DB 86 GAGCAAGCGGTGAGACAGAGCCGAGACCTGCGCAGACGACGAGCGAGAGC 145
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 DB 146 GGGCAGCGCTGGGAACATGCGACAGGCTTGTGGATTACTGGCGTGGGCGAGACA 205
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 DB 206 CTGCTGACAGAGTGCAGAGAGCTGCTCAGCTCCAGATCCAGCAAGTACGAGCGCG 265
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 |||||||
 DB 266 CTGATGACAGAGACCATGAAGGTTGAAGGCTTACAAATCGGAATCGAGCAACTG 325
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||||
 DB 326 ACCCGGCTGGGAGAGAGAGCGGCGAGCTGTCAAAGAGCTGACGCGCGCAGGCC 385
 QY 121 ArgLeuGlyAlaAspMetCyluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 DB 386 CCGCTGTGGCGCGGACATGAGAGACGTGTGCGCGCTGTGTCAGTACCGCGCGAGTGT 445
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluGluArgValArgLeuAlaSerHisLeuArg 160
 |||||||

DB 446 CAGGCCATGCTGGGCCAGACAGACCGAGAGACTGCGGCTGCGCTCCCACTGCC 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 506 AAGCTCGTAAAGAGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTAC 565
 QY 181 GlnAlaGlyAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 200
 |||||||
 DB 566 CAGCGCGGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 625
 QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln 220
 |||||||
 DB 626 CCCCCTGTGAGACAGAGTCCGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 685
 QY 220 OleuGlnGluArgAlaGlnAlaThrPheGlyGluArgLeuArgAlaArgMetGluGlu-Met 239
 |||||||
 DB 686 GTACAGAGAGCGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValArg 259
 |||||||
 DB 746 GGCAGCGCGGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 805
 QY 259 LysLeu---GluGluGlnAlaGlnGln-IleArgLeuGln-AlaGluAlaPheGlnAl 277
 ||:::
 DB 806 CCCAAGCTGAGACGAGCGAGCGCCAGCAAGATACGCTCAGAGCGCGAGCTTCTAGGC 865
 QY 277 ArgLeuLysSerTrpPheGluPro-----LeuValGluAspMetGlnArgGlnTrpAl 295
 |||:::|||||
 DB 866 CCGCGCTCAGAGAGTGTGTGAACCCCGTGGTTGGAACGACATAGCGCGCGCGCGAGTGG 925
 QY 295 ArgLeuValGluLysVal-GlnAlaAlaValGlyThrSerAlaAlaProValProSer 315
 |||:::|||||
 DB 926 CCGCTGTGTGAAAAAGTCCAGAGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCG 984
 QY 315 spasnhs 317
 |||||||
 DB 985 ACATATCAT 992

RESULT 2
 BG763371 842 bp mRNA linear EST 15-MAY-2001
 LOCUS 602735433F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4860585 5',
 DEFINITION
 mRNA sequence.
 ACCESSION BG763371 GI:14074024
 VERSION BG763371.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 842)
 NIH-MGC http://mgi.nci.nih.gov/.
 NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DT/PTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11718 row: b column: 10
 High quality sequence stop: 817.
 Location/Qualifiers

FEATURES
 source
 1..842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4860585"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOMB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 154 a 250 c 337 g 101 t

ALIGNMENT SCORES:

Alignment Scores:
Pred. No.: 7.46e-122 Length: 842
Score: 1254.00 Matches: 260
Percent Similarity: 97.04% Conservative: 2
Best Local Similarity: 96.30% Mismatches: 7
Query Match: 79.12% Indels: 2
DB: Gaps: 0

US-09-827-854-15 (1-317) x BG63371 (1-842)

OY 1 MetLVsValLeuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 36 ATGAAGGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 95
OY 21 GtUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpPheIns 40
DB 96 GAGCAAGCGGTGGAGACAGAGCGGAGCCGCGGCGCAGACAGACCGAGTGGAGAGC 155
OY 41 GtGlnArgTRPGLuLeuAlaLeuGlyArgPheTRPASPTrpLeuArgTRPValGlnThr 60
DB 156 GGCCAGCGCTGGAGACTGGCACTGGTCCCTTTGGGATCTGCGCGGTGCAGACA 215
OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 216 CTGTCTGAGCAGAGTCAGAGAGGAGTGCCTCAGCTCCAGAGTCCAGGAACTGAGAGCG 275
OY 81 LeuMetAspGluThrMetArgGluLeuArgGlySerGlySerGluGlnGluGlnLeu 100
DB 276 CTGATGGAGACACATGAGAGGTTGAGGCTTCAAAATCGAACTGGAGAGACAACTG 335
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerTrpGluGlnAlaAlaGlnAla 120
DB 336 ACCCGGCGGCGGAGAGAGCGGCGCGCTGCTCAAGAGAGCTGAGCGCGCGCAGCGC 395
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
DB 396 CGGCTGGGGCGGACATGAGAGAGCTGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 455
OY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 456 CAGGCCATGCTGGCCAGAGACACGAGAGAGTGGGCTGCGCTGCTGCCAAGCTGGCGC 515
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
DB 516 AAGCTGCGGTAAGCGGCTCTCCGCGATGCGATGACCTGCAAAAGCGCTGCGCACTGTA 575
OY 181 GlnAlaGlyAlaArgGluGlyValGluArgGlyLeuSerAlaTrpArgGluArgLeuGly 200
DB 576 CAGCGCGGGGCGCGGAGGCGCGCGCGCTGAGCGCCATCCGCGAGCGCTGGGGG 635
OY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyLysPro 220
DB 636 CCTTG-GTGGAGACAGGCGCGCTGGCGCGCACTGTGGCTGCTGGCGCGCAAGGTG 694
OY 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyLeuArgAlaArgMetGluLysMet-G1 240
DB 695 CTACAGAGAGCG 754
OY 240 ySerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaVal 260
DB 755 CAGCGGAGCCCGGAGCGCTGGAGAGGTGAAGGAGCAGAGTGGCGGAGTGCAGCAAG 814

OY 260 streuGluGlnAlaGlnGlnLeuArg 269
DB 815 CTGGGAGAGGAGCGCCAGCAGATACGC 842

RESULT 3
LOCUS B1668318 933 bp mRNA linear EST 12-SEP-2001
DEFINITION 603295681P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5', mRNA sequence.

ACCESSION B1668318
VERSION B1668318.1 GI:15582551

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 933)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNI at:

<http://image.llnl.gov>

Plate: L16M11798 row: a column: 21

High quality sequence stop: 780.

Location/Qualifiers

FEATURES

source

1..933

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5314844"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag)

; Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and

normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 171 a 284 c 359 g 119 t

ALIGNMENT SCORES:

Alignment Scores:
Pred. No.: 3.7e-121 Length: 933
Score: 1248.00 Matches: 263
Percent Similarity: 94.08% Conservative: 7
Best Local Similarity: 91.64% Mismatches: 14
Query Match: 78.74% Indels: 4
DB: Gaps: 1

US-09-827-854-15 (1-317) x B1668318 (1-933)

OY 1 MetLVsValLeuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
DB 75 ATGAAGGTTCTGTGGGCTGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGGTG 134
OY 21 GtUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpPheIns 40
DB 135 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCGCAGACAGACCGAGTGGAGAGC 194
OY 41 GtGlnArgTRPGLuLeuAlaLeuGlyArgPheTRPASPTrpLeuArgTRPValGlnThr 60
DB 755 CAGCGGAGCCCGGAGCGCTGGAGAGGTGAAGGAGCAGAGTGGCGGAGTGCAGCAAG 814

Db 195 GGCACGCGTGGAGACTGGCAGCTGGCTTTGGGATTACCTGCCCTGGGTGCAGACA 254
 QY 61 LeuSerGIuGIuValGIuGIuGIuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
 Db 255 CTGCTGAGCAGAGGTGAGGAGGAGCTGCTGAGCTCCAGGTCAACCGAGAACTGAGGGCG 314
 QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
 Db 315 CTGATGGACAGACCAAGAGGAGTTGAAAGCCTTACAAATCGCAACAGCGAGAAACAAC 374
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 Db 375 ACCCGCGTGGCGAGAGAGACCGGCGACGGCTGTCCAGAGAGCTGAGGGCGCGAGGCC 434
 QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGlnTyrArgGIuGIuVal 140
 Db 435 CGGCTGGGGCGGACATGAGAGAGCTGTGGCGGCTGTGTGACATGCCGGCGAGGTG 494
 QY 141 GlnAlaMetLeuGIuGIuInsThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 495 CAGCCATGCTCGCGCCAGACACCGAGAGCTCGGGTGGCTCCCTCCACCTGCGCG 554
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGIuLysArgLeuAlaValTyr 180
 Db 555 AAGCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 614
 QY 181 GlnAlaGIuAlaArgGIuGIuAlaArgGIuArgLysLeuSerAlaIleArgGIuArgLeuGIy 200
 Db 615 CAGGCGGGGGCGCGGAGGGCGCGAGCGGGCGCTCAGCCATCCGCGAGCGCCCTGGGG 674
 QY 201 ProLeuValGIuGIuGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuGIuPro 220
 Db 675 CCCCTGTGGAAACAGGCGCCGTCGGGGCGCCACCTGTGGGCTCCCTCGCGCGGCGAGCCG 734
 QY 221 LeuGIuGIuArgAlaGIuAlaIleArgLysGIuArgLeuArgAlaArg-MetGIu--GluMe 239
 Db 735 CTACAGAGAGGGGGCCAGGCTGTGGGCGAGCGGCTCGCGCGGATGTGAGGACACATTG 794
 QY 239 tGIySerArgTThrArgAspArgLeuAsp-GluValLysGIuGIuValAlaGIuValArg 259
 Db 795 GGGCAGCGGAGCCCGGAGAGCTGTGAAGAGAGGTGAAGAGAGCGAGGCGGAAGTGGGCG 854
 QY 259 IalysLeuGIuGIuGIuAlaGIuGIuIleArgLeuGIuAlaGIuAlaIleArgL 279
 Db 855 GCCAAAGCTGGGGGAGACAGGCCCGCAATTGCTTGACAGGCCAA-GCCTTCAGGCGCGGT 913
 QY 279 eulysSerTrrpHeGlu 284
 Db 914 CCAGAAGCTGGTTCANA 930
 RESULT 4
 LOCUS BG761746 938 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841411 5',
 mRNA sequence.
 ACCESSION BG761746
 VERSION BG761746.1 GI:14072399
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 938)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-riemail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LILCM1674 row: c column: 12
 High quality sequence stop: 767.
 Location/Qualifiers
 FEATURES
 source
 1..938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4841411"
 /clone_1ib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="PH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using 2AP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

BASE COUNT 171 a 272 c 374 g 121 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.32e-118 Length: 938
 Score: 1220.00 Matches: 272
 Percent Similarity: 92.62% Conservative: 4
 Best Local Similarity: 91.28% Mismatches: 11
 Query Match: 76.97% Indels: 11
 DB: 12 Gaps: 2
 US-09-827-854-15 (1-317) x BG761746 (1-938)
 QY 1 MetLysValLeuTrpAlaIleLeuValThrPheLeuAlaGIyCysGlnAlaLysVal 20
 Db 57 ATCAAGATTCTGTGGCTGTGCTGTGTCACATTCTTCGACAGAGATCCAGGCCAAGGTG 116
 QY 21 GIuGIuAlaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer 40
 Db 117 GAGCAAGCGGTGAGAGACAGAGCGCGAGCCGAGCTCGCAGACAGACCGAGTGGCAGAGC 176
 QY 41 GlyGlnArgTrpGIuLeuAlaLeuGIuArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db 177 GGCACAGCGCTGGGAACTGGCACAGGCTCTTTGGGATTACCTGGCGGCGGCGAGAA 236
 QY 61 LeuSerGIuGIuValGIuGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
 Db 237 CTGTCTGAGCAGAGTGCAGAGAGCTGTCTAGCTCCAGGTCAACCAGAACTGAGGGCG 296
 QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
 Db 297 CTGATGGACAGACCATGAAGAGGTGAAGGCGCTTACAAATCGGAACGTGAGGAAACATG 356
 QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
 Db 357 ACCCGCGTGGCGAGAGAGAGCGGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 416
 QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGlnTyrArgGIuGIuVal 140
 Db 417 CGGCTGGGGCGGACATGAGAGACGTGTCCGGCGCTGTGTCACTAGTACCGCGCGAGGTG 476
 QY 141 GlnAlaMetLeuGIuGIuInsThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 477 CAGGCCATGCTCGGCGCAGAGACCGGAGAGTGGGGTGGCTGGCTGCCACCTGCGCC 536
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGIuLysArgLeuAlaValTyr 180
 Db 537 AAGCTGCTTAAGCGGCTCTCCGCGATGACCTGCACAAAGGCGCTGGCAGGTGAC 596
 QY 181 GlnAlaGIuAlaArgGIuGIuAlaArgGIuArgLysLeuSerAlaIleArgGIuArgLeuGIy 200
 Db 597 CAGGCCGGGGCGCGAGGGCGCGAGCGCTGCTCAGCCATCGCGAGGCGCTGGGG 656

QY 201 ProLeuValIGluGlnGlyArgValAlaAlaThrValGlySer-LeuAlaGlyGlnPro 220
 DB 657 CCCCTGGTGGAAACAGAGCCGCGTGGGCGCCGCACTGTGGGTCTCCCTTGGCCGCGCACCC 716
 QY 220 cLeuGlnGluArgAlaGlnAlaATP-GlyGluArg-LeuArgAlaArgMetGlu---Glu 238
 DB 717 GCTACAGAGAGGCGCCAGAGCCCTGGGGGCGAGCGCTTGGCGCGCGGATGAGAGACAT 776
 QY 239 MetGlySerArgThr-Arg-AspArgLeuAspGluValLys-GluGlnValAla---Glu 256
 DB 777 TGGGACGCCGAGACCCGCGGACCGCTTGGAGAGGTAGAGGAGACAGTGGCGGAGG 836
 QY 257 ValArgAlaAlaLysGluGln-GlnAlaGlnGlnLeuArgLeuGlnAla---GluAlaPhe 275
 DB 837 TTGGGCGCCAGCTGAG 896
 QY 276 GlnAlaArgLeuLysSerTrpPheGlnProLeuVal 287
 DB 897 CCAGGCGGCTCAAAAGCTGGTTCGAGCCCTGGTT 932
 RESULT 5
 BI600906
 LOCUS 603249241F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',
 DEFINITION mRNA sequence.
 ACCESSION BI600906
 VERSION BI600906.1 GI:15493845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 942)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LLM11762 Row: a Column: 11
 High quality sequence stop: 762.
 Location/Qualifiers
 source
 1. 942
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5301010"
 /clone_id="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamH; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized for R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 177 a 284 c 361 g 120 t
 ORIGIN
 Alignment Scores: 6.13e-118 Length: 942
 Pred. No.: 1217.50 Matches: 266
 Score: 93.47% Conservative: 6
 Percent Similarity:

Best Local Similarity: 91.41% Mismatches: 13
 Query Match: 76.81% Indels: 8
 DB: 13 Gaps: 1
 US-09-827-854-15 (1-317) x BI600906 (1-942)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 73 ATGAAGGTTCTGTGGGCTGCGTGTGCTGATCATCTCCGGCAGATGTCAGGCGCAAGGTG 132
 QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 133 GAGCAACGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCAGCAGACAGCAGAGTGCAGAGC 192
 QY 41 GlyGlnArgTrpGluLeuAlaAlaLeuLysArgPheTrpAspTrpLysArgTrpAlaGlnThr 60
 DB 193 GGCACGCGCTGGGAACCTGCACTGGTGTGGTGGATTCACCTGCTGCTGGTGTGAGACA 252
 QY 61 LeuSerGluGlnValAlaGluGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 253 CTGTCTGACAGCTGACAGAGAGAGCTGCTCAGCTCCAGGTACCAAGAACTGAGGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnGlnLeu 100
 DB 313 CTGATGAGACGAGACCATGAGAGATTGAAGGCTTCAAAATCGGAACCTGGAGAGAACACTG 372
 QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 373 ACCCGGTGGCGAGAGACAGCGGCGACGCTGTCCAAGAGACTCAGAGCGCGCAGAGCC 432
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGlnVal 140
 DB 433 CGCGTGGCGCGGAGATGAGAGAGCTGTGGCGCCCTGTGTGACATACCGGCGCAGGTG 492
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 493 CAGGCGCATGCTCGCGCAGACAGCAGAGAGCTGCGGCTGCGCTCCACCTCGCGC 552
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspPheGlnGlnLysArgLeuAlaValTrp 180
 DB 553 AAGCTGCTTAAGCGGCTCTCCGCGATGCCGATGCTGCAAGAGCGCTGCACTGATGAC 612
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGlnArgGlyLeuSerAlaLysArgGluArgGly 200
 DB 613 CAGCGCGGCGCGCGGAGAGCGCGCAGCGCGGCTCAGCGCATCGGAGAGCGCTGGGG 672
 QY 201 ProLeuValIGluGlnGlyArgValAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 DB 673 CCCCTGGTGGAAACAGAGCGCGGCTGGCGGC-GCCACTGTGGGCTCCCTGCGCGCCAGCGC 731
 QY 221 LeuGlnGluArgAlaGlnAlaATPGLyGluArgLeuArgAla-ArgMetGlu-GluMetG 240
 DB 732 CTACAGAGAGCGGCGCCAGAGCTGCGGGCGAGCGGGGCGCGGCGAGATGAGAGCATTTG 791
 QY 240 LysSerArgThrArgAspArgLeuAsp-GluValLys-GluGlnValAlaGlnVal-ArgA 259
 DB 792 GCAGCGGAGACCGCGAGCGCTGAGACGAGAGGAGACCAAGTGGGGAAGTTGGCGCG 851
 QY 259 LysLysLeuGlnGlnGlnAla---GlnGlnLLeuArgLeuGlnAlaGlnAlaPheGlnAla 278
 DB 852 CCACCTTGGAGAACACAGCGCCAGCAGCAATACGCTTGCAGGCGCAGGCGCTTTCAGGCGC 911
 QY 278 rGluLeuLysSerTrpPheGlnPro 285
 DB 912 GCCTTCAAAAC-TGTTTGACCCCT 933
 RESULT 6
 BM042094
 LOCUS 603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
 DEFINITION mRNA sequence.
 ACCESSION BM042094
 VERSION BM042094.1 GI:16771361

KEYWORDS	EST.
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	1 (bases 1 to 800)
	NIH-MGC http://mgc.nci.nih.gov/ .
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: DCM/dmp
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 plate: LNCM1875 row: 1 column: 03
 High quality sequence stop: 792.

FEATURES	Location/Qualifiers
source	1. .800

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5420618"
/clone_1db="NIM_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="Dn10B (phage-resistant)"
/site_1="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCACGAG(c). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIM_MGC Library."

```

BASE COUNT	146 a	239 c	315 g	100 t
ORIGIN				

Alignment Scores:	
Pred. No.:	1,64e-116
Score:	1203.03
Percent Similarity:	99.20%
Best Local Similarity:	97.60%
Query Match:	75.90%
DB:	13
Length:	800
Matches:	244
Conservative:	4
Mismatches:	2
Indels:	1
Gaps:	0

US-09-827-854-15 (1-317) x BM042094 (1-800)

QY	1	MeuysValIeuTrpAlaIaleuLeuValThrPheueuIaGlyCysAlaIeVal	20
QY	50	ATGAAGGTCTGTGGGTGGGTGGTGTGCATCTCTGGCAGAGTCCAGGCCAAGGTG	109
QY	21	GIuInIaIaValGIuThrGIuProGIuProGIuIeuArgInGIuThrGIuTrpGIuSer	40
Db	110	GAGCAACGGGTGGAGACAGACCAGGACCCGAGCTGGCCACAGACAGACAGTGGCAGAC	159
QY	41	GIyInArGrTrpGIuIeuAlaIeuGIyArGPhetTrpAspTrIeuArGrTrpValGIuThr	60
Db	170	GGCCAGGGCTGGGAACCTGGCAGCTGGGTGGCTTTTGGGATTAACCTGGCGCTGGGTGCAGACA	229
QY	61	LeuSerGIuInValGIuInGIuIeuLeuSerSerGIuValThrGIuIeuArGAla	80
Db	230	CTGTCTGACGAGGTGCAGGAGGAGCTCTGAGTCCACGAGTCCACCGCAGAACTAGAGGGCG	289
QY	81	LeuMetAspGIuThrMetIeLysGIuIeuLysAlaTrIyLysSerGIuIeuGIuInGIuLeu	100
Db	290	CTGATGACGAGACCAAGAGGAAGATTGAAGCCCTACAAATCGGAACCTGGAGGAACAAC	349
QY	101	ThrProValIaGIuGIuThrArgAlaArgLeuSerLysGIuIeuGIuInIaIaIaGIuAla	120
Db	350	ACCCCGGTGGGAGGAGACGCGGGCAGCGCTGTCCAGAGAGCTGCACAGCGGCGCAGGCC	409

Qy	121	ArgLeuGluGlyAlaAspMetGluAspValcysGlyIleArgLeuValGlnIleArgGlyGluVal	140
Db	410	CGCGTGGGGCGCGAACAATGAGAGACGTGTGGCGCCCTGTGTCAATACCGCGGCAAGTG	469
Qy	141	GlnAlaMetLeuGlyGlnSerThrGluIleuLeuArgValIArgLeuAlaSerHisLeuArg	160
Db	470	CAGCCCATGTCTCGGCGCAAGCACCGAGAGAGCTGGCGGAGCCCTCTCCACCTGGGC	529
Qy	161	LysLeuArgIleValSerGluLeuLeuArgAspAlaAspAspleuGlnIleValSerArgLeuAlaValTyr	180
Db	530	AAGCTGCGTAAAGCCGGCTCTCCGGATGCCATGCACTGTCCAGAAAGCGCTGGCAATATAC	589
Qy	181	GlnAlaGlyAlaArgGlyGluGlyAlaGluIleuArgGlyLeuSerAlaIleArgGluIleuArgLeuGly	200
Db	590	CAGGCGCGGGGCGCGAGAGGCGCGAGCGGGCGCTCACAGCGCATCCGCGAGAGCGCTGGGG	649
Qy	201	ProLeuValGluGlnGlyArgValIArgAlaIaIaThrValGlySerLeuAlaGlyGlnPro	220
Db	650	CCCCGTGGTGAACAGGGGCGCGTGTGGGGCCAGCACTGTGGGCTCCCTGTGGCGGCGCAGCGG	709
Qy	221	LeuGlnGluArgAlaGlnAlaITrpGlyGluArgLeuArgAlaArgMetGluGlnMetGly	240
Db	710	CTAACAGAGCGCGGCCAACAGCCCTGGGGCCAGCGGTG-CGCGCGCGGATGAGAGATGGGCG	768
Qy	241	SerArgThrArgAspArgLeuAspGluVal	250
Db	769	ACGCGGAGCCGCAACGCTGTGACGAGCTG	798

RESULT	7
BG472299	
LOCUS	
DEFINITION	927 bp mRNA linear EST 21-MAR-2001 602513830F01 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:4645759 5', mRNA sequence.
ACCESSION	BG472299
VERSION	BG472299.1 GI:13404485

SOURCE ORGANISM	human.
	Homo sapiens

ORGANISM	Homo sapiens
	Chorata; Cranata; Vertebrata; Euteleostomi;
	Acyrota; Metazoa; Primates; Catarrhini; Homo.
REFERENCE	Mammalia; Eutheria; 1 (bases 1 to 927)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: L1CMI419 row: k column: 08
 Align quality sequence stop: 848.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_1ib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="hml0b (phage-resistant)"
/notes="Organ: eye; Vector: pOTB7; site_1: XhoI; site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCGCAGCAG(5'). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

```


Db 313 CTGATGACAGACCATGAGAGATTGAAGGCTTACAAATCGGAACCTGGAGAACACTG 372
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 373 ACCCGCGTGGCGAGAGACCGCGGACCGCTGTCCAAAGAGCTGCAGAGCGCGGAGGCC 432
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 433 CGGCTGGCGGCGACATGAGAGAGCTGTGGCGCGCTGTGTGACATACCGCGCGAGGTC 492
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 493 CAGGCCATCTCTGGCCAGAGACCGAGAGCTGCGGGTGGCTCCCTCCCTCCAGCTGCGC 552
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 553 AACCTCGTGAAGGCGCTCTCCCGATGCGGATGACCTGCAGAACGCGCTGCGCATGTAC 612
Qy 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 613 CAGCGCGGGGCGCGGAGGCGCGCGAGCGGCTC-AGCGCCATCCGCGAGCGCTGCGG 671
Qy 201 ProLeuVal-GluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGln-P 220
Db 672 CCCCTGGTGGAGACAGGCGGGGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 731
Qy 220 LeuGlnGluArg-AlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMet 239
Db 732 CGCTACAGAGCGGGGCGCGAGCTGGGGGAGCGGATACCGCGCGGATGAGACGACAG 791
Qy 240 GlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla 259
Db 792 GGGAGCGGACCGCGAGACCGCTGAGAGGTGAGAGACAGAGCGGGGCGGGGCGGCA 851
Qy 260 LysLeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluLaphegin 276
Db 852 CAAGCTGAAGAAACAGCCCAAGAAATACGCTTA--GCCGAGAGCTTCACAG 899
RESULT 9
Bg774871 817 bp mRNA linear EST 15-MAY-2001
LOCUS Bg774871 602649975F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4760827 5',
DEFINITION mRNA sequence.
ACCESSION Bg774871
VERSION Bg774871.1 GI:14045188
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LRCM1612 row: e column: 20
High quality sequence stop: 813.
Location/Qualifiers
1..817
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4760827"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

BASE COUNT 148 a 242 c 324 g 103 t
ORIGIN
Alignment Scores:
Pred. No.: 3.5e-115 length: 817
Score: 1190.50 Matches: 252
Percent Similarity: 97.32% Conservative: 2
Best Local Similarity: 96.55% Mismatches: 5
Query Match: 75.11% Indels: 4
DB: 12 Gaps: 0
US-09-827-854-15 (1-317) x Bg774871 (1-817)
Qy 1 MetLysValLeuThrPalaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 40 ATGAAGTTCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 99
Qy 21 GlnGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 100 GAGCAAGCGGTGGAGACAGACCGGAGCGGAGCTGCGCAGACAGACAGACAGAGAGAG 159
Qy 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 160 GGCACAGCGGTGGAACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219
Qy 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 220 CTGTCTGACAGAGTGCAG 279
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGluLeu 100
Db 280 CTGATGACAGACCATGAAAGAGGTGAAGGCTCAAAATCGGAATCGAGAGAACACTG 339
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluGlnAlaGlnAla 120
Db 340 ACCCGCGTGGCGAGAGACCGGGGACGCGCTGTCCAAAGAGCTGCAGCGCGGAGGCC 399
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 400 CGGCTGGCGGCGACATGAG 459
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 460 CAGGCCATCTCTGGCCAGAGACCGAGAGAGCTGCGGGTCCGCTGCCCTCCACCTGCGC 519
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 520 AAGCTCGTGAAGGCGCTCTCCCGATGCGGATGACCTGCAGAAAGGCTGGCAAGGTAC 579
Qy 181 GlnAlaGlyAlaArgGluGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 580 CAGCGCGGGGCGCGGAGGCGCGGAGCGGCGCTGAGCGCATCCGAGCGCGGCGGGG 639
Qy 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 640 CCCCTGTGAAGAGAGCGCGCTGCGGCGCTGACCTGCTGCTGCTGCTGCTGCTGCTG 698
Qy 221 LeuGlnGluArgAlaGlnAla--TrpGlyGluArgLeuArgAlaArgMetGluGluMetG 240
Db 699 CTACAGAGAGCGGGCCAGAGCCCTTGGGGGAGCGGATGCGCGCGGAGAGGAGATGG 758
Qy 240 LysSerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAla 259
Db 759 GCCACCGGAGCG--GAACCGCTGAGAGAGGTGAAGAGAGAGGTGGCGGAGGTGCGCGCC 815

RESULT 10
BI600563
LOCUS 603244936F1 NIH_MGC_96 811 bp mRNA linear EST 07-SEP-2001
DEFINITION mRNA sequence.
ACCESSION BI600563
VERSION BI600563.1 GI:15493502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM11726 row: 9 column: 10
High quality sequence stop: 783.
Location/Qualifiers
1..811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287329"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to 80T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGR), National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 148 a 245 c 316 g 102 t
ORIGIN
Alignment Scores:
Pred. No.: 4,99e-115 Length: 811
Score: 1189.00 Matches: 241
Percent Similarity: 98.37% Conservative: 1
Best Local Similarity: 97.97% Mismatches: 3
Query Match: 75.02% Indels: 1
DB: 13 Gaps: 0
US-09-827-854-15 (1-317) x BI600563 (1-811)
QY 1 MettysvalleuThrPALaAaleuValThrPhleuAlaGlyCysGlnAlaLysVal 20
DB 74 ATGAAGGTTCTGTGGCGTGGCTGTGCTGTCATCTCTGGAGATGCCAGGCCAAAGGTG 133
QY 21 GUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpIns 40
DB 134 GAGCAAGCGGTGGAGACAGCGGAGCCGAGCTGCCGACAGCAAGCGAGGAGAGC 193
QY 41 GlytlnArGTrpGluLeuAlaLeuGlyArGpHeTrPApTyLeuArGTrpValGlnTr 60
DB 194 GGCCAGCGGTGGAGACTGGCACTGGGTCTTTGGAGTACTCTGGCTGGGTGCAGACA 253
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 254 CTGCTTGAGCAGGTGCAGAGAGGCTGCTCACTCCAGGTCACCCAGGAAGTGAAGCGC 313

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGluLeu 100
DB 314 CTGATGAGCAGACACCATGATGAGAGTGAAGGCTCTCAATATCGAAGTGGAGACACACG 373
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAla 120
DB 374 ACCCGGTGGCGGAGAGACGCGGCGACGCTGTCCAAGAGACTGCAGCGCGCAGGCC 433
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyLys 140
DB 434 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCCCTGTGTCACTACCGCGGAGAGTG 493
QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 494 CAGGCGCATCTGGCCAGAGCACCAGAGAGCTGGCGGTGGCTCTCCACCTGGCC 553
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTr 180
DB 554 AAGCTGCTTAAGCGGCTCTCCGCGATGCCATGACCTGCACAAACGCTGGCAGTGTAC 613
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLysSerAlaIleArgGluArgLeu 200
DB 614 CAGGCGCGGCGCGGAGGCGCGGAGCGGCGCTACGCGCATCCGAGCGCGCTGGAGG 673
QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaIleArgLysSerLeuAlaGlyL 220
DB 674 CCCCTTGCTGAGAACAGGCGCGGCGGTGGCGCCGCCACTGTGGCTCTCCAGCAGCAGC 733
QY 220 OLeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMet 240
DB 734 GCTACAGAGAGCGGCGCCAGGCGCTGGCGGAGCGGCTGGCGCGCGGATGAGAGCAATGGG 793
QY 240 YSerArgThrArgAsp 245
DB 794 CAGCGGAAACCGCGAA 809
RESULT 11
BG829472
LOCUS 602763768F1 NIH_MGC_42 845 bp mRNA linear EST 22-MAY-2001
DEFINITION mRNA sequence.
ACCESSION BG829472
VERSION BG829472.1 GI:14177059
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 845)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LUCM1790 row: 9 column: 17
High quality sequence stop: 829.
Location/Qualifiers
1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4899112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;

Site-2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH-MGC Library. 1"

BASE COUNT 153 a 244 c 342 g 106 t
 ORIGIN

Alignment Scores:

Pred. No.:	1,4e-113	Length:	845
Score:	1175.50	Matches:	253
Percent Similarity:	96.23%	Conservative:	2
Best Local Similarity:	95.47%	Mismatches:	7
Query Match:	74.16%	Indels:	5
	12	Gaps:	2

US-09-827-854-15 (1-317) x BG829472 (1-845)

```

QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||
DB      57 ATGAAGTTCTGTGGGCTGTGCTGCTGTCACATTCTCGCAGATGCCAGGCCAAGTG 116
QY      21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
      |||||
DB      117 GACCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC 176
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||
DB      177 GGCCAGCGCTGGAGACCTGGACCTGGCTTTGGGATTACCTGGCTGGGTCGAGACA 236
QY      61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||
DB      237 CTCTCTGAGAGAGTGCAGAGAGAGCTGCTGCTCCAGGTCCACCCAGAGACTGAGGGCC 296
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
      |||||
DB      297 CTGATGAGACAGACCAATGAAGAGTTGAAGGCTTACAAATCGAGACTGAGAACTG 356
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
      |||||
DB      357 ACCCGGCTGGGAGAGAGAGCGGCGCACTGCTCCAAAGAGCTGAGCGCGCGCAGGCC 416
QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
      |||||
DB      417 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGGCTGGTGCAGTACCGCGCGCAGGTG 476
QY      141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      |||||
DB      477 CAGGCGATGCTCGGCCAGAGACCGAGAGCTCGGGTGGCGCTCCCTCCACCTGCGCG 536
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      |||||
DB      537 AACCTCGTAAAGGCGCTCCCGCGATGCCGATGACCTGAGAAAGCGCTGGGAGCTGTC 596
QY      181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      |||||
DB      597 CAGGCGGCGCGCGGAGAGCGCGGAGCGGCTTCAGCGCCATCCGCGAGCGCTGGGG 656
QY      201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGluGlnPro 220
      |||||
DB      657 CCCCTGCTGAACAGAGCGCGCGCGCGC -ACTGTGGGCTCCCTTGGCGGCAAG -CCG 714
QY      221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly 240
      |||||
DB      715 CTACAGAGAGCGGCGCGAGGTG ---GGGCGAGCGGCTCGCGCGGATGAGGAGATGGCG 771
QY      241 SerArgThr-ArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
      |||||
DB      772 AGCGGAGCGCGCA ---CGCTGGAGACAGGTGAAGAGACAGGTGGCGAGAGTGGCGCCAA 828
  
```

QY 260 sleuGluGluGln 264
 |||||
 DB 829 GCTTGAAGAGCAG 841

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 790)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihai
 Toshimiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM10672 row: e column: 21
 High quality sequence stop: 781.
 Location/Qualifiers
 1..790
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4793084"
 /clone_11b="NIH-MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="PH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTTTVN-3',
 size-selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."

FEATURES

BASE COUNT 143 a 234 c 311 g 102 t
 ORIGIN

Alignment Scores:

Pred. No.:	3.8e-111	Length:	790
Score:	1152.00	Matches:	236
Percent Similarity:	98.74%	Conservative:	0
Best Local Similarity:	98.74%	Mismatches:	3
Query Match:	72.68%	Indels:	1
	12	Gaps:	0

US-09-827-854-15 (1-317) x BG707147 (1-790)

```

QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||
DB      73 ATGAAGTTCTGTGGGCTGTGCTGCTGTCACATTCTCGCAGATGCCAGGCCAAGTG 132
QY      21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnTrpGlnSer 40
      |||||
DB      133 GAGCAAGCGGTGAGACAGACCGGAGCGGAGCTGCGCCACAGACCGAGTGGCAGAGC 192
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||
DB      193 GGCCAGCGCTGGAGACTGGACCTGGCTTTGGGATTACCTCGCTGGTGGCAGACA 252
  
```

OY	61	LeuSerGIuGlnValGIuGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuNArgIa	80
Db	253	CTGTGTAGACAGAGTGCAGAGAGAGTGTCTCAGCTCCACAGTCCACAGAACTGAGGCG	312
OY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu	100
Db	313	CTGATGGACGAGACCATTAAGAGACTTGAAGGCTTCMAATCGGACCTGGAGACAACTG	372
OY	101	ThrProValAlaGIuGIuThrArgIlaArgLeuSerLysGIuLeuGlnAlaIaGIuAla	120
Db	373	ACCCGGGTGGCGGAGAGAACCGGGCCAGCGCTGTCCAAAGAGACTCAGAGGGCGCAGGCC	432
OY	121	ArgLeuGIuAlaAspMetGIuAspAlaLysGIuArgLeuValGIuTyrArgIuVal	140
Db	433	CGGCTGGGCGCGGACATGTAGAGACCTGTGTGGCGCCCTGTGTACATACCGGGCCAGAGTG	492
OY	141	GlnAlaMetLeuGIuGlnSerThrGIuGIuLeuNArgValArgLeuAlaSerHisLeuNrg	160
Db	493	CAGGCATAGCTCTGGCCAAAGCACCCAGAGAGCTGCCGGTGTGGCTCTGCCACTTGGCC	552
OY	161	LysLeuNArgLysArgLeuLeuNArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	553	AAGCTGGCTAAGGGGCTCTCCGCATATCCATGTACCTGCAGAAAGCCCTGGCAATGTAC	612
OY	181	GlnAlaGIuAlaArgGIuGIuAlaGIuArgGIuLeuSerAlaIleArgGIuNArgLeuGIu	200
Db	613	CAGGCGGGGCGCCGCGAGAGCGCCAGCGGGTCTCACGCGCATCCGCGAGGCGCTGGGG	672
OY	201	ProLeuValGIuGIuGlnGIuArgValArgAlaAlaThrValGIuSerLeuAlaGIuPro	220
Db	673	CCCCGTGTGGAAACAGGCGCGGTGGGGCCGCCACTGTGTGGGTCTCCCTGGCAGCCAGCGG	732
OY	221	LeuGlnGIuArgAlaGlnAlaIleArgIleGIuNrgLeuNrgAlaArgMetHisGIuLeuMet	239
Db	733	CTACAGAGGCGGCGCCAGGCTGTGGGCGCAGCGGTG-CGCGCGCGGATGAGAGAGTGG	788

LOCUS	DEFINITION	ACCESSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
B1551475									
B1551475	603199.31.F1 NIH_MGC_95 mRNA sequence.	B1551475	B1551475.1 GI:15438787	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 919) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Miklos Palcovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shizaki Toshiki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov plate: LHAM1691 row: 1 column: 04 High quality sequence stop: 812.

```
FEATURES
source
location/Qualifiers
1. 919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_id="NHLMC_95"
/life_type="inprocampus"
/lab_host="DH10B"
```

```

BASE COUNT      172 a      270 c      363 g      114 t
ORIGIN
/Note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gcgcgaag
); Oligo-dn primed using primer 5'-TCTTTTCTTTTCTTTTCTTTT-3',
size-selected for average insert size 2.5 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHRI, National
Institutes of Health). Note: this is a NIH-MGC Library."

```

Alignment Scores:	
Pred. No.:	9.75e-11
Score:	1149.00
Percent Similarity:	93.63%
Best Local Similarity:	92.51%
Query Match:	72.49%
Ds:	13
Gaps:	0
Length:	919
Matches:	247
Conservative:	3
Mismatches:	16
Indels:	0

OY	1	MeltyVallelrrPALAlaleueneuAlThrPheuuAlagIcySGlAlaIaIyVal	20
OY	73	ATGAAGTTCGTGTGGGCTGGCTGTGGTCAATTCCTGGCAGATGCCAGGCCAAGGTC	13
OY	21	GIuGlnAlaIaValGIuThGcluProGluProGluLeuAArgGlnGlnThGcluTrpGlnSer	40
Db	133	GAGCAAGGGGTGTGACACAGAGGCCGAGCCAGACGTGCCACACAGACAGAGTGGCAGAGC	19
OY	41	GIyGlnAArgTPGluLeuAlaLeuGlyAArgPheTrpAspTrpLeuAArgTrpValGIuThr	60
Db	193	GGCCAGCCCTGGGAACTGGACACTGGGTGGCTTTGGATTACCTCGCTGGGTGCAGACA	25
OY	61	LeuSerGIuGlnValGlnGlnGluLeuLeuSerSerGlnValThThGlnGluLeuAArgAla	80
Db	253	CTGTCTGGACAGGTGCAGAGAGCTGTCTCACCTCCACAGTCCAGACAGAACTGAGGGCG	31
OY	81	LeuMetAspGIuThrMetGlyGIuLeuLysAlaTrpLysSerGIuLeuGlnGluGlnLeu	10
Db	313	CTGATGGACGACACCATGATGAGAGAGTTGAAGGCTTCAATTCGGAACCTGGAGAACAACTG	37
OY	101	ThrProValAlaGluGluThrAArgAlaAArgLeuSerLysGIuLeuGlnAlaIaIaGlnAla	12
Db	373	ACCCGGTGGCGGAGAGAACGCGGGCCAGCGCTCTCCAAAGAGACTCAGAGCGGCGCAGGCC	43
OY	121	ArgLeuGlnAlaIaAspMetGIuAspValLysGIuAArgLeuValGlnTrpAArgGlyValAl	14
Db	433	CGGCTGGGCGCGCATGTAGAGACGTGTGGGGCCCTGTGGTCACTACCCGCGCAGAGTG	49
OY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuAArgValAArgLeuAlaSerHisLeuAArg	16
Db	493	CAGGCCATGCTCGGCCAAGCACCGAAGAGCTGGCGGTGGCTCGCTCCACCTGGCGC	55
OY	161	LysLeuAArgLysAArgLeuLeuAArgAspAlaAspAspLeuGlnLysAArgLeuAlaValTrp	18
Db	553	AAGCTGCCTAAGCGGCTCTCCGCGATGCCGATGCACCTGACAGAAACGCCCTGGCAGCTTAC	61
OY	181	GlnAlaGlyAlaIaArgGIuGlyAlaIaGluAArgGlyLeuSerAlaIleAArgGluAArgLeuGly	20
Db	613	CAGCGGGGGGGCGGATGAGGCGGCCAGGCGGTGTCTACCGCCCATCCGCGAGGCGCTTGGGG	67
OY	201	ProLeuValGIuGlnGlyAArgValAArgValAlaAlaThrValIGlySerLeuAlaGlyGlnPro	22
Db	673	CCCCGTGTGGAAACAGAGGCCGCTGGCGGGCCGCCACTGTGGGCTCCTCGGGCGGACGCTT	73
OY	221	LeuGlnGlnAArgAlaGlnAlaIaTrpGlyGluAArgLeuAArgAlaAArgMetGluGluMetGlu	24
Db	733	A--CAGGAGCGGGGCGCAGGCTCGGGGCGAAGCGTCCGCGCGCGATGAGAGAAATAGG	79
OY	240	YSerAArgThrAArgAspAArgLeuAspGluValLysGIuGlnValAlaGluValAArgAlaIy	26
Db	791	CAGCGCGAGCCGCGGACG--CTGGACGAAGTGAAGAGCACGTGGCGGGAGGG--TGGCGCAA	84

QY 260 sLeuGlUGluGlnAlaGln 266
|||||:|||||
Db 848 GCTGAGAGACCAGCAGAG 866

RESULT 14
BM042228

LOCUS BM042228 757 bp mRNA linear EST 07-NOV-2001
DEFINITION 60361618F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 757)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2021 row: k column: 21
High quality sequence stop: 757.
Location/Qualifiers
1..757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5557004"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 134 a 229 c 236 g 98 t

ORIGIN

Alignment Scores:
Pred. No.: 8.38e-110 Length: 757
Score: 1139.00 Matches: 235
Percent Similarity: 99.16% Conservative: 0
Best Local Similarity: 99.16% Mismatches: 1
Query Match: 71.86% Indels: 2
Gaps: 0

US-09-827-854-15 (1-317) x BM042228 (1-757)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||:|||||
Db 50 ATGGAAGTTCTGTGGCTGCTGTGGTGCACATCTCTGCGAGGATGCGAGGCCAGGCTG 109

QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGluGlnThrGluTrpGlnSer 40
|||||:|||||
Db 110 GAGCAAGCGGTGAGACACAGCGGAGCGGAGCTCGCAGACAGCAGGCGCAGAC 169

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||:|||||
Db 170 GGCACAGCGCTGGAACTGGCAGCTGGCTTTGGGATTACCTGGCTGGGTGCAGACA 229

QY 61 LeuSerCluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||:|||||
Db 230 CTGCTGTGAGCAGGTGCGAGGAGAGCTGCTCAGCTCCAGCTCACAGGAACAGAGCGGC 289

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
|||||:|||||
Db 290 CTGATGAGCAGACCATGAAGAGGTTGMAAGCCTTACAAATCGGAACACTGAGGAACAAC 349

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||:|||||
Db 350 ACCCGGTGCGAGGAGAGACGCGGCGCTGTCCAGAGAGCTGCAGGCGGCGAGGCC 409

QY 121 ArgLeuGlnLysAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLysVal 140
|||||:|||||
Db 410 CGGCTGGCGCGGACATGAGGAGCGTGCGCGCGCTGTGCGACATCCGCGGAGG 469

QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuLysSerHisLeuArg 160
|||||:|||||
Db 470 CAGGCCATGCTCGGCCAGAGACCGAGAGACTCGGGGTGGCTCGCTCCACTGCGC 529

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||:|||||
Db 530 AACCTGCGTAAAGCGGCTCCTCCGATGCCGATGACCTGCAGAAAGCGCTGCGCAGTAC 589

QY 181 GlnAlaGlyAlaArgGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
|||||:|||||
Db 590 CAGCGCGGCG 649

QY 201 ProLeuValGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
|||||:|||||
Db 650 CCCCTGTGTAACAGAGGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 709

QY 220 OleuGlnGluArgAlaGlnAlaTrpGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 236
|||||:|||||
Db 710 GCTACAGAGACCGGCGC-CAGGCTGGGCGAGCGGCTGCGCGCGCGGATG 757

RESULT 15
BG706129

LOCUS BG706129 907 bp mRNA linear EST 07-MAY-2001
DEFINITION 60266903F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792030 5',
mRNA sequence.
ACCESSION BG706129
VERSION BG706129.1 GI:13981169
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 907)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10669 row: 1 column: 23
High quality sequence stop: 832.
Location/Qualifiers
1..907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4792030"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptPR (modified

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 1477.5 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALVLVFLAGCOAKV.....AREGAEGLSAIRRLGLPV 203

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO/spool/US09827854/rtnat_11032003_101610_27486/app_query.fasta_1.3576
-DB=genembl -QPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=humand40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09827854.gcgn.1.1_13008.etrnat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1014	100.0	1110	6 E00359	E00359 cDNA coding
2	1014	100.0	1110	6 E00823	E00823 DNA sequenc
3	1014	100.0	1147	6 AX302545	AX302545 Sequence
4	1014	100.0	1156	6 BD004278	BD004278 Apo E hum
5	1014	100.0	1156	6 HUMAPOE3	K00396 Homo sapien
6	1014	100.0	1186	6 BC003557	BC003557 Homo sapi
7	1007	99.3	1110	6 E08423	E08423 DNA coding
8	1006	99.2	1156	6 BD004277	BD004277 Apo E hum
9	1005	99.1	1157	6 AX333278	AX333278 Sequence
10	1005	99.1	1157	6 AX409597	AX409597 Sequence
11	1005	99.1	1157	6 I15975	I15975 Sequence 1
12	1005	99.1	1157	6 HUMAPOE	M12529 Human apoli
13	1002	98.8	1156	6 BD004279	BD004279 Apo E hum
14	984	97.0	660	6 A62340	A62340 Sequence 28
15	984	97.0	660	6 A62342	A62342 Sequence 30
16	984	97.0	660	6 A62344	A62344 Sequence 32
17	984	97.0	660	6 AR075563	AR075563 Sequence
18	920	90.7	1178	9 MFAP0E	X13887 Monkey mRNA
19	825.5	81.4	5491	9 AF261279	AF261279 Homo sapi
20	825.5	81.4	41907	6 AX358722	AX358722 Sequence
21	825.5	81.4	41907	9 AF050154	AF050154 Homo sapi
22	825.5	81.4	107567	9 AC011481	AC011481 Homo sapi
23	808	79.7	5515	9 HUMAPOE4	M10065 Human apoli
24	789	77.8	1138	4 AF303830	AF303830 Tupai a gi
25	782	77.1	5413	9 AF261280	AF261280 Pan trogl
26	769	75.8	1045	10 MUSAPOE	M12414 Mouse apoli
27	769	75.8	1104	10 BC028816	BC028816 Mus muscu
28	760.5	75.0	965	6 AX384545	AX384545 Sequence
29	760.5	75.0	1108	4 BRAPOLPE	X61171 B. taurus mR
30	760.5	75.0	5617	6 AX384541	AX384541 Sequence
31	760.5	75.0	6026	6 AX384539	AX384539 Sequence
32	749	73.9	208239	2 AC021988	AC021988 Homo sapi
33	742.5	73.2	1154	4 BRAPOMER	X64839 B. taurus mR
34	735.5	72.5	1060	4 RABAPOLP	M36603 Rabbit apol
35	730	72.0	4762	6 BABAPOL	M29322 Baboon apol
36	724.5	71.4	1126	6 ARI64342	ARI64342 Sequence
37	724.5	70.8	951	6 AR205885	AR205885 Sequence
38	718	70.8	951	6 S76779	S76779 RAP0E-apol
39	718	70.8	1069	10 RATAPOL	J00705 rat apoli
40	715.5	70.6	1122	4 S8APOL	X72835 S. scrofa mR
41	692	68.2	959	10 MUSAPOE	M73480 Mus musculu
42	675	66.6	528	9 H8APOE	X00199 Human mRNA
43	649	64.0	228698	2 AC127479	AC127479 Mus muscu
44	649	64.0	237653	2 AC073760	AC073760 Mus muscu
45	625	61.6	478	6 AX330507	AX330507 Sequence

RESULT 1

ALIGNMENTS

E00359			1110 bp	RNA	linear	PAT 29-SEP-1997
LOCUS	E00359	CDNA coding human apolipoprotein E3.				
DEFINITION	E00359					
ACCESSION	E00359.1	GI:2168646				
VERSION	JP 1985118189-A/L.					
KEYWORDS	Homo sapiens.					
SOURCE	Homo sapiens.					
ORGANISM	Homo sapiens.					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	Teratishii,Y., Tokimatsu,N., Matsui,Y., Kimura,M. and Ikeda,Y.					
JOURNAL	DNA FRAGMENT					
	Patent: JP 1985118189-A 1 25-JUN-1985;					
COMMENT	MITSUBISHI CHEM IND LTD					
	OS human					
	PN JP 1985118189-A/1					
	PD 25-JUN-1985					
	PF 29-NOV-1983 JP 1983224980					
	PI TERANISHII YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI					
	KIMURA MASAKO,					
	PI IKEDA YASUKO					
	PC C12N15/00,C07H21/04//C12P21/00;					
	CC strandedness: Double;					
	CC topology: Linear;					
	CC hypothetical: No;					
	CC anti-sense: No;					
	CC *source: tissue_type=liver;					
	FH key Location/Qualifiers					
	FT CDS 15..968					
	FT /Product='apolipoprotein E3 precursor' FT					
	FT sig_peptide 15..68					
	FT /Product='apolipoprotein E3 signal peptide' FT					
	FT mat_peptide 69..965					
	/product='apolipoprotein E3'					
FEATURES						
source	Location/Qualifiers					
	1..1110					
	/organism='Homo sapiens'					
	/db_xref='taxon:9606'					
BASE COUNT	198 a 353 c 416 g 143 t					
ORIGIN						
	Alignment Scores:					
	Pred. No.: 1.53e-71					
	Score: 1014.00					
	Percent Similarity: 100.00%					
	Best Local Similarity: 100.00%					
	Query Match: 100.00%					
	Gaps: 0					
DB:						
US-09-827-854-15_COPY_1_203 (1-203) x E00359 (1-1110)						
OY 1 MeltyValLeuTTPPALAlaleuleuValThrPhelAleAlAGlycysGlnAlalysVal 20						
Db 15 ATGAAGGTTCGTGGCGCTGCCTGCTGCATTCCGGAGAATGCCAGGCCAAAGTG 74						
OY 21 GlnGlnAlaValAlGluThrGlupProglupProgluleuArygInngInnThrGlutTrpGlnSer 40						
Db 75 GAGCAAGCGGGTGGACAACAGACCAGCCCAAGCTCGCCCAAGCACAGAACCGAGTGGCAGAGC 134						
OY 41 GlyGlnArgrTRPGluleAlaleuclYarPhetRpaSPTRYleuArgrTVaJglnThr 60						
Db 135 GGCCAGCGCCGTGGAACTGCACACTGGGTGGCTTTGGGATTAACCTGCGCTGGGTGCAGACA 194						
OY 61 LeuserrGluGlnValGlnGlnGluLenuLeuserrserGlnValThrGlnGluLeuArgala 80						
Db 195 CTGTCTGACGAGGTGCAGAGAGAGCTGCTCATCCCAAGGTCCACCAAGAACGTAGGGCG 254						
OY 81 LeulleAspgLuThrMetTySGluLeuLysAlaTryLysserGluLeugluGlnGlnLeu 100						
Db 255 CTGATGACAGCAACCATTAAGAGATTGAAGGCTCTTCAATAATGGAACTGGAGAGAACTGG 314						

QY	101	ThpRoVAlaIaIuGluGluThrArgAlaAArgLeuSerLySGluuEngInlaIaIaGlnAla	120
Db	315	ACCCGCGTGGCGGAGAGACCGCGGCGCTGTCCAAAGAGCTGCAGCGCGCGACGCC	374
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrArgIArgIyluVal	140
Db	375	CGGCTGGGCGCGAGACATGAGAGACGTGTGGCGCGCTGGTGCAGTACCGCGCGAGG	434
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCATGCTCGCGCCAGACACCGAGAGACTCGGGGTCTGCCTCCCTCCACCTCGCG	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	495	AAGCGCTAAGCGCGCTCTCCGCGATGCCGATGACCTCGAGAGCGCTGCGACGTGAC	554
QY	181	GlnAlaGlyAlaAArgGluGlyAlaGluArgGlyLeuSerAlaIleArgIuArgLeuGly	200
Db	555	CAGCGCGGCGCGCGCGAGAGGCGCGAGCGCGGCTCAGCGCCATCCGCGAGGCGCTGGGG	614
QY	201	ProLeuVal 203	
Db	615	CCCCGTGG 623	
RESULT 2			
LOCUS	E00823	1110 bp	DNA linear PAT 29-SEP-1992
DEFINITION	DNA sequence coding for human apolipoprotein E and its signal peptide.		
ACCESSION	E00823		
VERSION	E00823.1	GI:2169084	
KEYWORDS	JP 1986096997-A/1.		
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 1110)		
AUTHORS	Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.		
TITLE	PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN		
JOURNAL	Patent: JP 1986096997-A 1 15-MAY-1986;		
COMMENT	MITSUBISHI CHEM IND LTD		
	OS Human (Homo sapiens)		
	PN JP 1986096997-A/1		
	PD 15-MAY-1986		
	PF 16-OCT-1984 JP 1984216987		
	PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUOKO, KIMURA MASAKO		
	PC C12P21/00,A61K35/74,A61K37/04,C12N15/00,(C12P21/00,C12R1.19),,		
	PC (C12N15/00;		
	PC C12R1.19);		
	CC strandedness: Double;		
FEATURES	CC topology: Linear;		
	CC hypothetical: No;		
	CC anti-sense: No;		
	*source: tissue= Liver;		
	Key Location/Qualifiers		
	FT	3'UTR	1..14
	FT	sig_peptide	15..68
	FT	/product='human apolipoprotein E signal peptide'	
	FT	CDS	69..968
	FT	/product='human apolipoprotein E' FT	3'UTR
969..1110.			
Location/Qualifiers			
1..1110			
/organism="unidentified"			
/db_xref="taxon:32644"			
BASE COUNT			
198 a 353 c 416 g 143 t			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.53e-71 Length: 1110			
Score: 1014.00 Matches: 203			
Percent Similarity: 100.00% Conservative: 0			

Best Local Similarity: 100.00%
Query Match: 100.00%
Mismatches: 0
Indels: 0
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x E00823 (1-1110)

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
15 ATGAAGGTTCTGTGGGCTGCGCTTGGTGCATCTCTGGCAGATGCCAGATGCCAAGGTC 74
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
75 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACAGAGGAGAGC 134
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
135 GGGCAGCGCTGGAGACTGGCACTGGCTCTTTGGATTAACCTGCGTGGTGCAGACA 194
OY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
195 CTGCTGAGCAGCTCAGAGAGAGCTGCTCAGCTCCAGAGTCACCCAGAACTGAGGGCG 254
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
255 CTGATGAGCAGACCATGAAAGAGTTGAAGGCTTCAAAATCGAACTGGAGAGCAACTG 314
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
315 ACCCGGTGGCGGAGAGAGCGGGCGAGCTGCTCCAAAGAGCTCAGAGCGGCGAGGCC 374
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
375 CGGCGTGGCGCGAGACATGAGAGAGCTGCGCGCCCTGCTCAGTACCGCGGCGAGGTC 434
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
435 CAGGCCATGCTCGGCGAGAGCACCAGAGAGCTGCGGCTGCGCTGCCCATCTGGCGC 494
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
495 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACTGTCAGAAAGCGCTGGCAATGTAC 554
OY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
555 CAGCGCGGGGCGCGGAGGCGCGCAGCGGCGCTCAGCGCATTCGGAGAGCGCTGGGG 614
OY 201 ProLeuVal 203
615 CCCCTGGTG 623
Db

RESULT 3
AX302545
LOCUS AX302545
DEFINITION Sequence 63 from Patent WO01/5177.
ACCESSION AX302545
VERSION AX302545.1 GI:17383082
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Morin, P.J., Sherman-Bauslt, C.A., Pizer, E.S. and Hough, C.D.
TITLE Tumor markers in ovarian cancer
JOURNAL Patent: WO 01/5177-A 63 11-OCT-2001.
THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES
Location/Qualifiers
1. 1147
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 210 a 365 c 425 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1,59e-71 Length: 1147
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AX302545 (1-1147)

OY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
46 ATGAAGGTTCTGTGGGCTGCGCTTGGTGCATCTCTGGCAGATGCCAGAGGTC 105
OY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
106 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTGCGCAGACAGACAGAGGAGAGC 165
OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
166 GGGCAGCGCTGGAGACTGGCACTGGCTCTTTGGATTAACCTGCGTGGTGCAGACA 225
OY 61 LeuSerGluGlnValGlnGlnLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
226 CTGCTGAGCAGCTCAGAGAGAGCTGCTCAGCTCCAGAGTCACCCAGAACTGAGGGCG 285
OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
286 CTGATGAGCAGACCATGAAAGAGTTGAAGGCTTCAAAATCGAACTGGAGAGCAACTG 345
OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
346 ACCCGGTGGCGGAGAGAGCGGGCGAGCTGCTCCAAAGAGCTCAGAGCGGCGAGGCC 405
OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
406 CGGCGTGGCGCGAGACATGAGAGAGCTGCGCGCCCTGCTCAGTACCGCGGCGAGGTC 465
OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
466 CAGGCCATGCTCGGCGAGAGCACCAGAGAGCTGCGGCTGCGCTGCCCATCTGGCGC 525
OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
526 AAGCTGCGTAAGCGGCTCTCCGCGATGCGCATGACTGTCAGAAAGCGCTGGCAATGTAC 585
OY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
586 CAGCGCGGGGCGCGGAGGCGCGCAGCGGCGCTCAGCGCATTCGGAGAGCGCTGGGG 645
OY 201 ProLeuVal 203
646 CCCCTGGTG 654
Db

RESULT 4
BD004278
LOCUS BD004278
DEFINITION Apo E humanized mammal.
ACCESSION BD004278
VERSION BD004278.1 GI:18632239
KEYWORDS JP 2001017028-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Fujita, S., Hamanaka, H., Fukui, Y. and Yokoyama, M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 2 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
COMMENT OS Homo sapiens (human)
PN JP 2001017028-A/2
PD 23-JAN-2001

QY	141	GIN1A1MettLeuGLYChInSerThrcInuLeuAryValAryLeuAlaSerfistLeuArg	160
Db	481	CAGGCATGCTGGCGACAGCACCGAGGAGCTGGCGGGTGGCTGGCCCTCCACCTGGCG	540
QY	161	LysLeuAryGLysAryGLeuLeuAryAspAlaAspAspLeuGLNLYsAryGLeuAlaValTyr	180
Db	541	AAGCTGGCTAAGCGGCTCTCCCGGCGATCCGATGACCTGCAGAAAGCGGCTGGCACTGTAC	600
QY	181	GIN1A1AGLYAlaAryGLuGLYAlaGLuAryGLYLeuSerAlaIleAryGLuAryLeuGLY	200
Db	601	CAGGCGGGGGCCCCGAGAGGCGCCGAGGCGCGCTCAGAGCGGCATCCGAGGCGCTGGGG	660
QY	201	ProLeuVal 203	
Db	661	CCCCGTGTC 669	
RESULT 6			
LOCUS	BC003557	1186 bp	mRNA
DEFINITION			1186 bp mRNA linear PRI 12-JUL-2001
ACCESSION	BC003557		
VERSION	BC003557.1		
KEYWORDS	MGc.		
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			

[illegible]


```

Db 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCAGACAGAGCGAGAGC 181
Qy 41 GtGlnArqTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArqTrpValGlnThr 60
Db 182 GGCACAGCGCTGGAGACTGCGACTGGTGGTGGATTTGAGTACCTCGCTGGGTGAGACA 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArqAla 80
Db 242 CTGTCTGAGACAGGTGACAGAGAGCTGCTCAGCTCCCAAGTACCCAGAACTGAGAGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
Db 302 CTGATGAGACAGACCATGAGAGAGTTGAAGCGCTTCAAAATCGGAATCGAGAGACACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArqLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAAGCGGAGAGAGCGGCGACGCTGTCACAGAGACTGACAGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgLysGluVal 140
Db 422 CGGCTGGGCGGCGACATGAGAGAGCGTGGCGCCGCTGTGTCAGTACCGCGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArqValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTCGCGCAGAGACACCGAGAGCTGCGGCTGCGCTCCACCTGGCGC 541
Qy 161 LysLeuArqLysArgLeuLeuArqAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAACGGGCTCTCCGCGATCCCGATGACCTGCAAGACGCGCTGGACGTGAC 601
Qy 181 GlnAlaGlyAlaArqGluGlnGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 602 CAGGCGCGGCGCGCGGAGGCGCGCCAGCGCGGCTCAGCGCCATCCGCGAGGCGCTGGGG 661
Qy 201 ProLeuVal 203
Db 662 CCCCTGGTG 670

```

RESULT 11

```

LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5472858.
ACCESSION 115975
VERSION 115975.1 GI:1250883
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1157)
AUTHORS Attie,A.D., Gretsch,D.G., Shurley,S.L. and Beckage,N.E.
TITLE Production of recombinant proteins in insect larvae
JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
FEATURES
source
1..1157
/organism="unknown"

```

BASE COUNT 212 a 370 c 426 g 149 t

ALIGNMENT SCORES:

Pred. No.: 8,24e-71 Length: 1157
 Score: 1005.00 Matches: 201
 Percent Similarity: 99.01% Conservative: 0
 Best Local Similarity: 99.01% Mismatches: 2
 Query Match: 99.11% Indels: 0
 DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x 115975 (1-1157)

```

Qy 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGGTTTCTGTGGCTGCGCTGTCACATTCCTGGCAGATGCGAGGCAAGGTG 121

```

```

Qy 21 GtGlnAlaValAlaGluThrGluProGluProGluLeuArqGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCAGACAGAGCGAGAGC 181
Qy 41 GtGlnArqTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArqTrpValGlnThr 60
Db 182 GGCACAGCGCTGGAGACTGCGACTGGTGGTGGATTTGAGTACCTCGCTGGGTGAGACA 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArqAla 80
Db 242 CTGTCTGAGACAGGTGACAGAGAGCTGCTCAGCTCCCAAGTACCCAGAACTGAGAGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
Db 302 CTGATGAGACAGACCATGAGAGAGTTGAAGCGCTTCAAAATCGGAATCGAGAGACACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArqLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCGGTAAGCGGAGAGAGCGGCGACGCTGTCACAGAGACTGACAGCGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgLysGluVal 140
Db 422 CGGCTGGGCGGCGACATGAGAGAGCGTGGCGCCGCTGTGTCAGTACCGCGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArqValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTCGCGCAGAGACACCGAGAGCTGCGGCTGCGCTCCACCTGGCGC 541
Qy 161 LysLeuArqLysArgLeuLeuArqAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAACGGGCTCTCCGCGATCCCGATGACCTGCAAGACGCGCTGGACGTGAC 601
Qy 181 GlnAlaGlyAlaArqGluGlnGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgLeuGly 200
Db 602 CAGGCGCGGCGCGCGGAGGCGCGCCAGCGCGGCTCAGCGCCATCCGCGAGGCGCTGGGG 661
Qy 201 ProLeuVal 203
Db 662 CCCCTGGTG 670

```

RESULT 12

```

LOCUS HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
DEFINITION Human apolipoprotein E mRNA, complete cds.
ACCESSION M12529
VERSION M12529.1 GI:178848
KEYWORDS
SOURCE Homo sapiens (clone: PHAE[112,178,813].) male 57-year old liver
CDNA to mRNA.
ORGANISM Homo sapiens

```

REFERENCE 1 (bases 1 to 1157)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 McLean,J.W., Elshoubagy,N.A., Chang,D.J., Mahley,R.W. and
 Taylor,J.M.
 Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing
 of a new variant

J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source
 1..1157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE[112,178,813]."
 /sex="male"
 /tissue="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 /gene="APOE"

```

CDS
/note="apoE mRNA: G00-119-691"
62..1015
/gene="APOE"
/note="precursor"
/codon_start=1
/product="apolipoprotein E"
/protein_id="AAB59518.1"
/db_xref="GI:178849"
/db_xref="GDB:G00-119-691"
/translation="MKVMAALLVPLFACCAKVEQAVETPEPEELRQOTEMOGGWR
ELALGFMDLYRWQTLSEVOEELISVQTELRALMETKELKAYSPLEQTP
VAETFRARLSELOTAQARLQADMEQVGRVQYRGEVQAMIGQSTVEELRVLASLR
KLRRRLRDPDDIQRLAVQAGARGARGLSALRELRPLVEQGRVAAVATGSLAG
QPLQERAAQMGSERLRAMEMGSRTRDRLEQVKEQVAVRKLIEQAOQIRLQAEAFQ
ARKSWFEPLVEDMOROMAGLVEKVAAGVTSAAVPVSDNH"
62..115
/gene="APOE"
/note="G00-119-691"
116..1012
/gene="APOE"
/product="apolipoprotein E"
/note="G00-119-691"
370
/gene="APOE"
/note="a in PHAE[112,178]: g in PHAE813; G00-119-691"
/replace="g"
410
/gene="APOE"
/note="a in PHAE[112,178]: g in PHAE813; G00-119-691"
/replace="g"
569
/gene="APOE"
/note="c in PHAE[112,178]: g in PHAE813; G00-119-691"
/replace="g"
784
/gene="APOE"
/note="t in PHAE[112,178]: c in PHAE813; G00-119-691"
/replace="c"
859
/gene="APOE"
/note="g in PHAE[112,178]: a in PHAE813; G00-119-691"
/replace="a"
variation
BASE COUNT 212 a 426 g 149 t
ORIGIN 30 bp upstream of BstNI site.

Alignment Scores:
Pred. No.: 8,24e-71 length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x HUMAPOE (1-1157)
QY 1 MetlyValleuTrpAlaalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 62 ATGAAGGTCTGTGGGTGCTGTGCATCTCTGCGAGAGATGCCAGGCCAAGGTC 121
|||||
QY 21 GUGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTyrGlnSer 40
|||||
DB 122 GAGCAAGCGGTGAGACAGAGCGGAGCCGACCTGGCGCAGACAGCCGAGTGGCAGAGC 181
|||||
QY 41 GUGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
DB 182 GGGCAGCGCTGGGACTGGCTGGCTTTGGGATTACTGCGCTGGTGGCGAGACA 241
|||||
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
DB 242 CTGCTGTGAGAGAGTGCAGAGAGTGTGCTCACTCCCAAGTCAACCAAGACTGAGGGCG 301
|||||
QY 81 LeuMetAspGluThrMetlyGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||

```

```

DB 302 CTGATGACGACGACCATGAAGAGTTGAGGCGCTACAAATCGGAACTGGAGAACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLySGluLeuGlnAlaAlaGlnAla 120
|||||
DB 362 ACCCGGTACCGAGGAGAGAGCGGCGGCGCTCTCCAGAGAGTGCAGACGGCGAGGCC 421
|||||
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||
DB 422 CGGCTGGGCGCGGACATGAGGAGACGTGTGGCGCGCGCTGTGCATACCGCGCGAGGTG 481
|||||
QY 141 GlnAlaMetLeuGlnGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
DB 482 CAGCGCATGCTGGCCAGACACCGAGAGAGCTGGGTGGCTCCCTCCACCTGGCGG 541
|||||
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
DB 542 AACCTGCTAAGGCGCTCTCCGCGCATCCGATGACCTGAGAAAGCGCTGGCAGTGC 601
|||||
QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
DB 602 CAGCGCGGCGCGCGGAGGCGCGGAGCGCGCTCAGCGCCATCCGCGCGAGCGCTGGG 661
|||||
QY 201 ProLeuVal 203
DB 662 CCCCTGGTG 670

RESULT 13
BD004279 1156 bp DNA linear PAT 31-JAN-2002
LOCUS BD004279
DEFINITION Apo E humanized mammal.
ACCESSION BD004279
VERSION BD004279.1 GI:18632240
KEYWORDS JP 2001017028-A/3.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1156)
AUTHORS Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.
TITLE Apo E humanized mammal
JOURNAL Patent: JP 2001017028-A 3 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)
PN JP 2001017028-A/3
PD 23-JAN-2001
PF 28-APR-2000 JP 2000128919
PR
PI SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA PC
A01K67/027,A61K45/00,A61P25/28,A61P43/00,C12N5/10, PC
C12N15/09//C07K14/775,
PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC

FEATURES
source FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..1156
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 208 a 369 c 432 g 147 t
ORIGIN

Alignment Scores:
Pred. No.: 1,42e-70 length: 1156
Score: 1002.00 Matches: 202
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 98.82% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BD004279 (1-1156)
QY 1 MetlyValleuTrpAlaalaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

```

Db 61 ATGAAGTTCTGTGGGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGTGTG 120
QY 21 GtuglnAlaValaGlutThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
Db 121 GAGCAAGCGGTGGAGACACAGCCGGAGCCGAGCTGCGCCAGCAGACCAAGTGGCAGAGC 180
QY 41 GtlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 181 GGCACGCTGGGAATGTCACATGGGTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValaGlnGlnGluLeuLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
Db 241 CTGTCTGACAGCTGACAGAGAGAGCTGCTCACCAGTCCACAGCAGACTGAGAGCGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
Db 301 CTGATGAGACGACACCATGAGAGCTTGAAGGCTTACAAATCCGAACTGGAGACACTG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCCGCTGGCGAGAGAGACGCGGCGACGCTGTCCAAGAGACTGACAGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValaGlnTyrArgGlyGlnVal 140
Db 421 CGGCTGGGCGGACATGAGAGACGTGCGCGCTGTGTCAGTACCGCGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnValGlnSerThrGlnGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATCTCGCGCCAGAGACACCGAGAGCTGCGGGTCCGCTCGCTCCCACTGCGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGGCTCTCCGCGATGCGCATGACCTGCAAGAGCGCTGGCAGCTGAC 600
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 601 CAGGCGGGGCGCGGAGGCGCGGCGCTCAGCGCATCCGCGAGCGCGCTGGGG 660
QY 201 ProLeuVal 203
Db 661 CCCCTGGTG 669
RESULT 14
LOCUS A62340 660 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 28 from Patent WO9712992.
ACCESSION A62340
VERSION A62340.1 GI:3716293
KEYWORDS
SOURCE
ORGANISM
unidentified.
unidentified.
unclassified.
REFERENCE 1 (bases 1 to 660)
AUTHORS Van, L. F., Burbach, J. P. and Grosveld, F. G.
TITLE DIAGNOSIS METHOD AND REAGENTS
JOURNAL Patent: WO 9712992-A 28 10-APR-1997;
ROYAL NETHERLANDS ACADEMY OF A (NL);
Other publication AU 7142796 970428.
COMMENT
FEATURES
location/Qualifiers
1. 660
/organism="unidentified"
/db_xref="taxon:32644"
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03641.1"
/db_xref="GI:3716294"
/translation="PSGEGRSPGADPITGRKMKVLAALVTPLAGCOAKVEOAV
ETPEPELROOTEMOGORWELAIGRMDYLRWOTLSSEVOEELSSOYQELRALM
DETMEKELKAYKSELBOLTPVAETFRALSELTOAORLADMEDVDGCRVYQGEY
QAMLGOSTEELRVLASHLRKLRRLRDLDPDLOKRLAVTQAGAREGERSAISIRE
LG"
BASE COUNT 129 a 197 c 247 g 87 t

ORIGIN
Alignment Scores:
Pred. No.: 2e-69 Length: 660
Score: 984.00 Matches: 197
Percent Similarity: 98.99% Conservatve: 0
Best Local Similarity: 98.99% Mismatches: 2
Query Match: 97.04% Indels: 0
DB: 6 Gaps: 0
US-09-827-854-15_COPY_1_203 (1-203) x A62340 (1-660)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAAGTTCTGTGGGCTGCTGTGTCACATTCCTGGCAGATGCCAGGCAAGTGTG 121
QY 21 GtuglnAlaValaGlutThrGluProGluProGluLeuArgGlnGlnThrGluProGlnSer 40
Db 122 GAGCAAGCGGTGGAGACACAGCCGGAGCCGAGCTGCGCCAGCAGACCAAGTGGCAGAGC 181
QY 41 GtlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCACGCTGGGAATGTCACATGGGTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValaGlnGlnGluLeuLeuLeuSerSerGlnValaThrGlnGluLeuArgAla 80
Db 242 CTGTCTGACAGCTGACAGAGAGAGCTGCTCACCAGTCCACAGCAGACTGAGAGCGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
Db 302 CTGATGAGACGACACCATGAGAGCTTGAAGGCTTACAAATCCGAACTGGAGACACTG 361
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 362 ACCCCGCTGGCGAGAGAGACGCGGCGACGCTGTCCAAGAGACTGACAGCGCGCAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValaGlnTyrArgGlyGlnVal 140
Db 422 CGGCTGGGCGGACATGAGAGACGTGCGCGCTGTGTCAGTACCGCGCGAGGTG 481
QY 141 GlnAlaMetLeuGlnValGlnSerThrGlnGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATCTCGCGCCAGAGACACCGAGAGCTGCGGGTCCGCTCGCTCCCACTGCGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAGCGGGCTCTCCGCGATGCGCATGACCTGCAAGAGCGCTGGCAGCTGAC 601
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 199
Db 602 CAGGCGGGGCGCGGAGGCGCGGCGCTCAGCGCATCCGCGAGCGCGCTGGCGCGCTG 658
RESULT 15
LOCUS A62342 660 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 30 from Patent WO9712992.
ACCESSION A62342
VERSION A62342.1 GI:3716295
KEYWORDS
SOURCE
ORGANISM
unidentified.
unidentified.
unclassified.
REFERENCE 1 (bases 1 to 660)
AUTHORS Van, L. F., Burbach, J. P. and Grosveld, F. G.
TITLE DIAGNOSIS METHOD AND REAGENTS
JOURNAL Patent: WO 9712992-A 30 10-APR-1997;
ROYAL NETHERLANDS ACADEMY OF A (NL);
Other publication AU 7142796 970428.
COMMENT
FEATURES
location/Qualifiers
1. 660
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 129 a 197 c 247 g 87 t

ORIGIN

Alignment Scores:

Pred. No.:	2e-69	length:	660
Score:	984.00	Matches:	197
Percent Similarity:	98.99%	Conservative:	0
Best local Similarity:	98.99%	Mismatches:	2
Query Match:	97.04%	Indels:	0
DB:	6	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x A62342 (1-660)

QY	1	MethylsalLeuTrpAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaIleVal	20
Db	62	ATGAAGGTTCTGTGGCTGCGTTCCTGTGTACATTCTGGCAGAGATGCCAGCCAAAGTG	121
QY	21	GIuGlnAlaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer	40
Db	122	GAGCAAGCGGTGAGACAGAACCCGAGCTCCGACAGACAGACCGAGTGGCAGAGC	181
QY	41	GIyGlnArgTrpGluLeuAlaIleuGlyArgPheTrpAspIlyrLeuArgTrpValGlnThr	60
Db	182	GGCCAGCGCTGGGAACCTGGCACTGGTGGCTTTGGGATTAACCTGGCGCTGGGTGCAGACA	241
QY	61	LeuSerGIuGlnValGlnGIuGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla	80
Db	242	CTGTCTAGGACAGGTGCAGAGAGAGACTGCTAGCTCCCAAGTACCCAGAACTAGAGGCG	301
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaIleTrpLysSerGIuLeuGlnGIuGlnLeu	100
Db	302	CTGATGGACGAGACCATGAAGAGATTGAAGGCTCAAAATCGGAACGTGAGAAACAATC	361
QY	101	ThrProValAlaGIuGlnThrArgAlaArgLeuSerLysGIuLeuGlnAlaIleGlnAla	120
Db	362	ACCCCGTAGCGGAGGAGACGCGGGCACGGCTGTCCAAAGAGAGCTGCAGACGGCGCAGGCC	421
QY	121	ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGlnTrpArgGlyGIuVal	140
Db	422	CGGCTGGGCGGGACATGAGAGAGCTGTGGGCGCGCTGGGTGGCAGTACCGCGCGAGGTG	481
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCCAATGCTCGGCCAGAGACCGGAGAGACTGGGGTGGCGCTGCCACTGGCGC	541
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	542	AAGCTGGCTAAGCGGCTCTCCCGGATCCCGATGACTCTGCAGAAAGCGCTCGCAGAGTATC	601
QY	181	GlnAlaGlyAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeu	199
Db	602	CAGGCTGGGGGCCCCGAGAGGGGCGGACGGGGCTTACGGCCCAATCGCGAGAGCCCTTG	658

Search completed: March 14, 2003, 17:32:22
Job time : 1481.5 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 114.454 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALVTFLAGCOAKV.....AREGAEGRGLSAIRERLGPLV 203

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-O=/cgn2.1/USPRO_spool/US09827854/runat_11032003_101609_27476/app_query.fasta.1.3576
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=PTC -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=200000000
-USER=US09827854.cgn.1.1.1201 @runat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGQUERY_NES_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: Geneseq_101002:*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1014	100.0	954	24	AAD26035	Human apolipoprotein E
2	1014	100.0	1110	7	AAN60409	Human apolipoprotein E
3	1014	100.0	1147	22	ABA83113	Human apolipoprotein E
4	1014	100.0	1156	22	AAF84315	Human apolipoprotein E
5	1014	100.0	1156	24	AAD22048	Human apolipoprotein E
6	1014	100.0	1279	22	AAS22437	Human apolipoprotein E
7	1010	99.6	1156	24	AAD22052	Human apolipoprotein E
8	1006	99.2	1156	22	AAF84314	Human apolipoprotein E
9	1006	99.2	1156	24	AAD22049	Human apolipoprotein E
10	1006	99.2	1156	24	AAD22051	Human apolipoprotein E
11	1005	99.1	1157	17	AT06957	Human apolipoprotein E
12	1005	99.1	1157	24	ABN95746	Human apolipoprotein E
13	1005	99.1	1157	24	ABK64514	Human apolipoprotein E
14	1005	99.1	1157	24	ABL65450	Human apolipoprotein E
15	1004	99.0	1110	6	AAN50450	Human apolipoprotein E
16	1002	98.8	1156	22	AAF84316	Human apolipoprotein E
17	1002	98.8	1156	24	AAD22047	Human apolipoprotein E
18	999	98.5	1156	24	AAD22050	Human apolipoprotein E
19	984	97.0	660	18	AAE69792	Human apolipoprotein E
20	891.5	87.9	1107	19	AAE75756	Human apolipoprotein E
21	825.5	81.4	9360	24	AB131915	Human apolipoprotein E
22	825.5	81.4	10716	24	AAD26034	Human apolipoprotein E
23	820	80.9	3805	20	AAZ09524	Human apolipoprotein E
24	820	80.9	3805	20	AAZ09526	Human apolipoprotein E
25	803.5	79.2	10716	24	AAD26108	Human apolipoprotein E
26	760.5	75.0	965	24	AAD32081	Human apolipoprotein E
27	760.5	75.0	5617	24	AAD32077	Human apolipoprotein E
28	760.5	75.0	6026	24	AAD32075	Human apolipoprotein E
29	756	74.6	1381	22	AAS22673	Human apolipoprotein E
30	724.5	71.4	1126	19	AAV29159	Human apolipoprotein E
31	678	66.9	610	20	AAE98955	Human apolipoprotein E
32	651	64.2	407	24	ABK34238	Human apolipoprotein E
33	625	61.6	478	24	ABN95924	Human apolipoprotein E
34	625	61.6	478	24	ABL62679	Human apolipoprotein E
35	625	61.6	478	24	ABL67340	Human apolipoprotein E
36	625	61.6	478	24	ABL67340	Human apolipoprotein E
37	559	55.1	936	17	AAE18070	Human apolipoprotein E
38	559	55.1	936	17	AAE18070	Human apolipoprotein E
39	521	51.4	597	12	AAQ11980	Human apolipoprotein E
40	514	50.7	330	12	AAQ11980	Human apolipoprotein E
41	512	50.5	597	15	AAQ69099	Human apolipoprotein E
42	504	49.7	405	21	AAE02139	Human apolipoprotein E
43	485	47.8	345	22	AAH98479	Human apolipoprotein E
44	419	41.3	260	21	AAA40342	Human apolipoprotein E
45	410	40.4	382	21	AAA44191	Human apolipoprotein E

ALIGNMENTS

RESULT 1	
AAD26035	standard; cDNA, 954 BP.
XX	
XX	AAD26035;
XX	26-MAR-2002 (first entry)
XX	
XX	Human apolipoprotein E (APOE) cDNA.
XX	Human; anti-lipemic; neuroprotective; nontoxic; genetic variant; APOE;
XX	apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
XX	genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
XX	atherosclerosis; single nucleotide polymorphism; chromosome 19q13.2; ss.
OS	Homo sapiens.
XX	
XX	Key
XX	Location/Qualifiers

FT	CDS	1..954	/*tag- a	/product= "human APOE protein"	
FT			replace (13, C)		
FT	variation		/*tag- a		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (31, G)		
FT			/*tag- b		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (364, A)		
FT			/*tag- c		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (388, C)		
FT			/*tag- d		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (498, G)		
FT			/*tag- e		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (526, T)		
FT			/*tag- f		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
FT	variation		replace (622, A)		
FT			/*tag- g		
FT			/standard_name=	"Single nucleotide polymorphism (SNP) "	
XX					
PN	MO200179234 -A2.				

US-09-827-854-15_copy_1_203 (1-203) x AAD26035 (1-954)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGluCysGlnAlaIleVal 20

Db 1 ArgAAAGGTTCTGTGGGCTGGCTTCTGTGCATTTCTCGAGGATGCCAGGCCAAAGTG 60

QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuAArgGlnGlnThrGluTrpGlnSer 40

Db 61 GAGCAAGAGGGGTGGAGACAGACGCCGAGCCGAGCTGGCGCAGCAGACGAGTGCGCAAGC 120

QY 41 GlyGlnArgTrpGluLeuAlaAlaLeuGlyIleArgPheTrpAspPylIleuArgTrpValGlnThr 60

Db 121 GCCCGGCCCTGGGAAGCTGGCACTGGGTGCTTTGGGGTTAACTCGCCCTGGGTGCACACA 180

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80

Db 181 CTGTTGTGACGAGGTGCAGGAGAGACTGCTACAGCTCCCAAGGTACCCAGACCAACTGAGGGCG 240

QY 81 LeuMetAspGluThrMetLysGluLeuLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100

Db 241 CTGATGGACGAGACCATGAAGAGATTGAAGGCCCTACAAATTCGCAACTGGAGGAACAACCTG 300

QY 101 ThrProValAlaGluGluThrArgAlaArguSerGluGluGlnAlaAlaGlnAla 120

Db 301 ACCCGGGTGGGAGAGAACCGCGGCAAGGCTGTCTCAAGAGAGCTGCAGGGGGCGCAGGCC 360

PE 16-APR-2001; 2001WO-US12303.
 XX
 PR 14-APR-2000; 2000US-197188P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 XX
 PI Choi JY, Klem SE, Koshy B, Lee HH;
 DR WPI: 2002-075064/10.
 XX P-PSDB: AAE15158.
 PT Genotyping human apolipoprotein gene of individual for determining
 PT haplotype of individual, involves determining identity of nucleotide
 XX pair at specific polymorphic sites for two copies of gene -
 PS
 PS Claim 26; Fig 2; 78pp; English.
 XX
 CC The patent discloses novel genetic variants of human apolipoprotein
 CC E (APOE) gene. The invention also relates to compositions and methods
 CC for haplotyping and/or genotyping the APOE gene. The haplotyping
 CC methods of the invention are useful for improving the efficacy and
 CC reliability of several steps in the discovery and development of
 CC drugs for treating diseases associated with APOE activity, e.g.,
 CC familial dysbetalipoproteinemia, type III hyperlipoproteinemia,
 CC atherosclerosis, and Alzheimer's disease. They are useful to validate
 CC APOE as a candidate agent for treating a specific condition or disease
 CC predicted to be associated with APOE activity and in the design of
 CC clinical trials of candidate drugs for treating a specific condition
 CC or disease predicted to be associated with APOE activity. Genotyping
 CC or haplotyping methods are useful to screen for compounds targeting
 CC APOE to treat a specific condition or disease associated with APOE
 CC activity. The present sequence is a cDNA encoding human APOE protein.
 CC APOE gene is located on chromosome 19q13.2.
 XX
 XX Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;

Alignment Scores:	
Pred. No.:	1.18e-87
Score:	1014.008
Percent Similarity:	100.008
Best Local Similarity:	100.008
Query Match:	100.008
DB:	24
Length:	954
Matches:	203
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

ID	AAN60409 standard; DNA; 1110 BP.	
XX		
XX	AAN60409;	
AC		
XX		
D7	01-JAN-1980 (first entry)	
XX		
DE	Human apolipoprotein-E.	
XX		
KW	Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	15..968
FT		/*tag= a
XX		
PN	AU8547513-A.	
XX		
PD	24-APR-1986.	
XX		
PF	17-SEP-1985;	85AU-0047513.
XX		
PR	11-JUN-1985;	85JP-0126989.
XX		
PA	(MITU) MITSUBISHI CHEM IND KK.	

XX WP1: 1986-150217/24.
 DR P-PSDB: AAP60507.
 XX New DNA sequence coding for human apolipoprotein-E - and
 PT expression vectors and transformed cells contg. it
 XX
 XX Disclousure; Fig 2; 45pp; English.
 XX
 CC The encoded protein is used to treat subjects who are deficient in
 CC apolipoprotein-E (or who produce abnormal forms of this molecule)
 CC and therefore are likely to suffer from hyperlipidemia, resulting in
 CC arteriosclerosis. It can also be used to raise antisera for
 CC detecting the protein deficiency or production of abnormal forms.
 XX
 SQ Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other;

Alignment Scores:
 Pred. No.: 1.42e-87 Length: 1110
 Score: 1014.00 Matches: 203
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 7 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AAN60409 (1-1110)

OY 1 MetLysValLeuTPAlaAlaLeuLeuValThrpheLeuAlaGlyCysGlnAlaVal 20
 DB 15 ATGAAGGTTCTGTGGCGTGGCTGTGTCACATTCTCGGAGGATGCCAGGCCAAGGTG 74
 OY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 75 GAGCAGAGGGTGGAGACAGACGCCGAGCGCGCCAGCAGACGACGAGTGGCAGAC 134
 OY 41 GlyIleuArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
 DB 135 GGCCAGCCCTGGAGACTGGCAGCTGTGGGATTCACCTGCGCTGGTCAGACACA 194
 OY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 195 CTGCTGAGCAGGTGCAGAGAGAGAGCTCTCAGCTCCAGGTCACCCAGGAACAGAGGCG 254
 OY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
 DB 255 CTGATGGACGAGACCATGATGAGAGCTTGAAGCGCTCAATATCGAAGCTGGAGAACAACTG 314
 OY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 315 ACCCGGTGGGAGAGACACGGGGCAGGCTGTCCAAAGGAGCTGCAGGCGCGCAGGCGC 374
 OY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
 DB 375 CGGCTGGGCGCCGACATGAGAGAGCTGTGGCGCCCTGCTGTCAGTACCGGCGAGGTG 434
 OY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 435 CAGGCGATGTCGGCGCAGACGACGAGAGCTGGCGGGTGGCGCTGCCACTCCAGCTGGCGC 494
 OY 161 LysLeuAlaGlyLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 DB 495 AAGCTGCTAAGCGGCTCTCCCGCATGCCATGACCTGCAGAAAGCCGCTGGCAGTGTAC 554
 OY 181 GlnAlaGlyAlaArgGluGlyValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 555 CAGCGCGGGGCGCCGAGAGGCGCCGAGCGGCTCAGCGCCATTCGAGAGCGCTCGGGG 614
 OY 201 ProLeuVal 203
 DB 615 CCCCTGGTGG 623

RESULT 3
 ABA83113

ID ABA83113 standard; DNA: 1147 BP.
 XX
 AC ABA83113:
 XX
 DT 08-FEB-2002 (first entry)
 XX
 DE Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.
 XX
 KW Ovarian tumour marker gene; human; overexpression; upregulation;
 KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;
 KW identification; serous cystadenoma; borderline serous tumour;
 KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;
 KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;
 KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;
 KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;
 KW immune response pathway; cell proliferation regulation; protein folding;
 KW membrane localised; secreted; therapeutic target; cytosolic;
 KW gene therapy; vaccine; ds.
 KW
 KW Homo sapiens.
 OS
 OS WO200175177-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 03-APR-2001; 2001WO-US10947.
 XX
 XX 03-APR-2000; 2000US-194336P.
 XX
 XX (US\$) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;
 DR WP1: 2001-626450/72.
 DR P-PSDB: ABB50287.
 XX
 PT Detecting and identifying ovarian tumor, identifying increased risk for
 PT developing ovarian cancer, and determining effectiveness of ovarian
 PT cancer treatment, by measuring expression level of ovarian tumor marker
 PT gene -
 XX
 PS Claim 23; Page 105-106; 140pp; English.
 XX
 XX The invention relates to methods for diagnosing and prognosing ovarian
 XX tumours in an individual via the detection and measurement of the
 XX expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,
 XX ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,
 XX ABA83181 and ABA83183). The methods of the invention are useful for
 XX detecting an ovarian tumour in a patient, for identifying an individual
 XX at increased risk for developing ovarian cancer, in prognostic tests for
 XX assessing the relative severity of ovarian cancer, in tests for
 XX monitoring a patient in remission from ovarian cancer and in tests for
 XX monitoring disease status in a patient being treated for ovarian cancer.
 XX The methods can additionally be used to identify a particular tumour as
 XX being an ovarian tumour (i.e., an epithelial ovarian tumour selected from
 XX serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,
 XX mucinous cystadenoma, borderline mucinous tumour, mucinous
 XX cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,
 XX clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner
 XX tumour. The ovarian tumour marker genes of the invention were identified
 XX using SAGE (serial analysis of gene expression) and were found to be
 XX overexpressed in a broad variety of ovarian epithelial tumour cells
 XX relative to normal ovarian epithelial cells. The marker genes are
 XX implicated in immune response pathways, in the regulation of cell
 XX proliferation and in protein folding, and many of these are membrane-
 XX localised or secreted. In addition to their use as diagnostic and
 XX prognostic markers, the ovarian tumour marker genes or their encoded
 XX proteins may be used as therapeutic targets for the treatment and
 XX prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,
 XX ABA83182 and ABA83184 represent the ovarian tumour marker genes of
 XX the invention.
 XX
 SQ Sequence 1147 BP; 210 A; 365 C; 425 G; 147 T; 0 other;

QY	181	GlnAaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaAlaArgGluArgGluGly	200
Db	601	CAGGCGGGGCGCCGAGGAGGCGGACGCGGCGCTCAGGCCATCGCGAGCGCCTGGG	660
QY	201	ProLeuVal 203	
Db	661	CCCCGCTG 669	
RESULT 5			
AAD22048			
ID	AAD22048	standard; DNA; 1156 BP.	
XX	AAD22048;		
XX	12-FEB-2002	(first entry)	
XX			
DE	Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.		
KM	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;		
XX	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	61..1014	
FT		/*tag= a	
FT	sig_peptide	/product= "Human apoE isoprotein, apoE3"	
FT		61..114	
FT		/*tag= b	
FT	mat_peptide	115..1011	
FT		/*tag= c	
FT		/product= "Mature human apoE isoprotein, apoE3"	
XX			
XX			
XX			
PD	18-OCT-2001.		
XX			
PF	06-APR-2001; 2001MO-US11358.		
XX			
PR	06-APR-2000; 2000US-0544386.		
PR	04-OCT-2000; 2000US-0679088.		
PR	05-APR-2001; 2001US-0827854.		
XX			
XX	(KOSP-) KOS PHARM INC.		
PA	(UYBO-) UNIV BOSTON.		
XX			
PI	Zannis VI, Kypreos KE;		
XX			
DR	WPI; 2002-010885/01.		
DR	P-PSDB; AAE13294.		
XX			
PT	New apolipoprotein E polypeptide and nucleic acid, useful for lowering		
PT	cholesterol, delaying the onset of or treating atherosclerosis in		
PT	mammal, without inducing hypertriglyceridaemia		
XX			
PS	Claim 14; Page 81; 91pp; English.		
XX			
CC	The present sequence is a human apolipoprotein E (apoE)		
CC	isoprotein, apoE3 DNA. The apoE lipoproteins are useful for		
CC	lowering cholesterol, delaying the onset of atherosclerosis,		
CC	treating or regressing atherosclerosis without inducing		
CC	hypertriglyceridaemia, in a mammal lacking an endogenous,		
CC	normally functioning apoE gene for low density lipoprotein (LDL)		
CC	receptor or is at risk for developing atherosclerosis due to		
CC	accumulation of lipoprotein remnants in the bloodstream or having		
CC	a defect in remnant removal.		
XX			
XX			
SQ	Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;		
Alignment Scores:			
Pred. NO.:	1,49e-87	Length:	1156
Score:	1014.00	Matches:	203

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x AMD22048 (1-1156)

QY	1	MellysValLeuTrpPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
Db	61	ATGAAGCTTCTGTGGCGCTGCTGCTGCTGCATATTCCTCGCAGATGCGCAGGCCAAAGGTG	120
QY	21	GLUGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnIleThrGluTrpGlnSer	40
Db	121	GAGCAAGGGGTGGAGACAGACCGGACCTGGCTGGCCAGCGACGACCAAGTGGCAGAC	180
QY	41	GLYGlnAQTTPGulLeuAlaLeuGlyAArgPheTrpAspTyrLeuArgTrpValGlnThr	60
Db	181	GGCCAGCGCTGGGAACGTGGCACTGGCTGCTTTTGGGATTACTCGCTGGGTGGCAGACA	240
QY	61	LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla	80
Db	241	CTGTGTGACGAGGTGGAGAGAGACTGCTCAGCTCCAGGTCCACCGAAGACTGAGAGGCG	300
QY	81	LeuMetAspGluTrpMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu	100
Db	301	CTGATGGACGAGACCATGAAAGAGTGTAAAGCCCTACAAATCGGAACCTGAGAGAACACTG	360
QY	101	ThrProValAlaGluGluTrpArgAlaAArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCCGGTGGCGGAGAGACACCGGGCGCTGTCAGAGAGCTGCAGGCGGCGCAGGCC	420
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyAArgLeuValGlnTyrArgGlyGluVal	140
Db	421	CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGGTGCAGTACCGCGGCGAGGTG	480
QY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuAlaArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCATGCTCGGCCAGAGACCGAGAGACTCGGGTGGCGCTGCCCTCCACCACTGGCGC	540
QY	161	LysLeuAArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	541	AAGCTGCGCTAAGCGGCTCTCTCCGGATGCCGATAGCTTCAGACAGCGCTGGCAGTGTAC	600
QY	181	GlnAlaGlyAlaAArgGluGluGlyAlaGluAArgGlyLeuSerAlaIleArgGluAArgLeuGly	200
Db	601	CAGCGGGGGCGCGCGAGGGGGCGGAGCGGGCTCAGCGCCATCCGCGAGCGCCCTGGGG	660
QY	201	ProLeuVal 203	
Db	661	CCCCGTGTG 669	

RESULT 6
AAS22437
ID AAS22437 standard; cDNA: 1279 BP.

AAS22437:
24-Oct-2001 (first entry)

Human cDNA encoding a novel human protein #3.

Human; novel protein; ss; Antinaemic; osteopathic; antiinflammatory; immunomodulatory; cytosatic; neuroprotective; vulnerability; neotopic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antisthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.

XX Homo sapiens.
XX
XX
PN W0200155437-A2.
XX

PD 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US02623.
 XX 25-JAN-2000; 2000US-0491404.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI
 XX
 XX
 DR MPI: 2001-451939/48.
 DR P-PSDB; AA014132.
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,
 PS nervous system disorders, and for regenerating bone and cartilage -
 Claim 1: Page 167-169; 894pp: English.

CC The invention relates to polynucleotides encoding novel human
 CC proteins or their active domains. The polypeptides, polynucleotides and
 CC antibodies raised against the polypeptides are used in a method of
 CC treatment of a mammal and prevention of disorders caused by the aberrant
 CC protein expression or activity. The polypeptides can be used as
 CC molecular weight markers, food supplements, and in antibody production.
 CC The polypeptides are used to identify compounds which bind to the
 CC polypeptides. Polynucleotides of the invention are used as probes and
 CC primers, for sequencing, for chromosome or gene mapping, in the
 CC production of recombinant proteins, and in generating anti-sense DNA or
 CC RNA and in gene therapy. Polypeptides of the invention can be used to
 CC target drugs to a tumor, in assays to determine biological activity, to
 CC raise antibodies/ elicit an immune response, to determine quantitative
 CC protein levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence encodes a protein of the invention.

XX
 XX
 SQ Sequence 1279 BP; 238 A; 402 C; 476 G; 163 T; 0 other;

Alignment Scores:
 Pred. No.: 1,68e-87 Length: 1279
 Score: 1014.00 Matches: 203
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x AAS22437 (1-1279)

QY 1 MetIysValleuTrpAlaalaLeuValTrhPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 159 ATGAAGTTCTGTGGCTGTGGCTGTGCATCTCTGGCAGAGAGCCAGGCCAAGGTG 218
 QY 21 GluGlnAlaValaGlnThrGluProGluProGluLeuArgGlnGlnThrGluPglIns 40
 DB 219 GAGCAAGCGGTGGAGACAGAGCCGAGCCGACCTGCGCAGACAGCCGAGTGGCAGAGC 278
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 279 GGCACAGCGCTGGAACTGGCACTGGGTCTTTGGGATTACTCTGGCGTGGGTCAGACA 338
 QY 61 LeuSerGlnGlnAlaGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 339 CTGTCTGAGAGAGGTGACAGAGAGTGTCTACGCTCCACAGTACCCAGCAACTGAGGGCG 398
 QY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100

DB 399 CTGATGAGCAGACCAATGAAGGAGTTGAMGGCCCTACAAAATCGCAACTGGAGAACAACTG 458
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 DB 459 ACCCGGTGGCGAGGAGAGCGCGGCAAGGCTGTCCAAAGAGCTGCAAGCGGCGAGGCC 518
 QY 121 ArgLeuGlnAlaAspMetGlnaAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 DB 519 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGTGACGTACCGCGCGAGGTG 578
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 579 CAGCCCATGCTCGCGCAGAGACACCGAGAGCTCGGGGTGCGCTCCCTCCACCTGCGC 638
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 639 AACCTGCGTAAAGCGGCTCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 698
 QY 181 GlnAlaGlnAlaArgGlnGlnGlnAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 DB 699 CAGCGCGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 758
 QY 201 ProLeuVal 203
 DB 759 CCCCTGTGTG 767

RESULT 7
 AAD22052
 ID AAD22052 standard; DNA; 1156 BP.
 XX AAD22052;
 AC
 XX
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
 XX
 XX Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 61..1014
 FT /tag= a
 FT /product= "Human apoE allele, apoE2**"
 FT sig_peptide 61..114
 FT /tag= b
 FT mat_peptide 115..1011
 FT /tag= c
 FT /product= "Mature human apoE allele, apoE2**"
 PN WO200177136-A1.
 PD 18-OCT-2001.
 XX
 XX
 XX 06-APR-2001; 2001WO-US11358.
 PF
 XX
 XX 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 XX (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 PA
 PI Zannis VI, Kyriacs KE;
 XX
 XX
 DR MPI: 2002-010885/01.
 DR P-PSDB; AA013298.
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -

|||||
Db 182 GGCAGCGCTGGGAACGACGACTGGCTTTGGATTACCTGGCTGGTGACAGACA 241
Qy 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuAArgAla 80
Db 242 CTGTCTAGAGAGGTGCGAGGAGAGCTGCTCAGCTCCCAAGTCAACCCAAAGACGAGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 302 CTGATGACAGACACATGAAGAGCTTAAGGCTTACAAATCGGAACTGGAGGACAACTG 361
Qy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGAGGAGAGCGGCGGCTGTCGCAAGAGCTGCAGAGCGGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGGCTGGTGAGTACCGCGGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlnSerThrGlnGlnLeuAArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGGCGGACAGACCGAGAGCTGGGGTGCGCTCGCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGGCTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 602 CAGGCGGGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661
Qy 201 ProLeuVal 203
Db 662 CCCCTGGTG 670
RESULT 13
ABK64514
ID ABK64514 standard; DNA: 1157 BP.
AC ABK64514;
DT 18-JUN-2002 (first entry)
XX Human benign prostatic hyperplasia gene #409.
DE Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
KM Homo sapiens.
OS Homo sapiens.
PN WO200212440-A2.
XX 14-FEB-2002.
PD 07-AUG-2001; 2001WO-US24708.
PE 07-AUG-2000; 2000US-223323P.
XX 05-JUN-2001; 2001US-0873319.
PA (GENE-) GENE LOGIC INC.
XX (NISR) JAPAN TOBACCO INC.
PI Munger WE, Kulikarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
PS Disclosure; Page 239-240; 444pp; English.
XX The invention relates to a method of diagnosing (I) the onset or

CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
Alignment Scores:
Pred. No.: 1,07e-86 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.018 Conservative: 0
Best Local Similarity: 99.018 Mismatches: 2
Query Match: 99.11% Indels: 0
DB: Gaps: 0
US-09-827-854-15_copy_1_203 (1-203) x ABK64514 (1-1157)
Qy 1 MetLysValLeuThrPalaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 62 ATGAGAGTTCTGTGGGCTGTGGCTGTGGTGCACATTCCTGGCAGAGATCCAGGCAAGGTG 121
Qy 21 GlnGlnAlaValGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db 122 GAGCAAGCGGTGAGACAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 181
Qy 41 GlyGlnArgTrpLysLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db 182 GGCAGCGCTGGAGACTGGACACTGGGCTGTTGGGATTACCTGGCTGGTGACAGACA 241
Qy 61 LeuSerGluGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
Db 242 CTGTCTAGAGAGGTGCGAGGAGAGCTGCTCAGCTCCCAAGTCAACCCAAAGACGAGGCG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
Db 302 CTGATGACAGACACATGAAGAGCTTAAGGCTTACAAATCGGAACTGGAGGACAACTG 361
Qy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTAGCGAGGAGAGCGGCGGCTGTCGCAAGAGCTGCAGAGCGGCGAGGCC 421
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 422 CGGCTGGGCGGACATGAGGAGAGCTGTGGCGGCTGGTGAGTACCGCGGCGAGGTG 481
Qy 141 GlnAlaMetLeuGlnSerThrGlnGlnLeuAArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCCATGCTGGCGGACAGACCGAGAGCTGGGGTGCGCTCGCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 542 AAGCTGGCTAAGCGGCTCTCCGCGATCCCGATGACCTGCAGAAAGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlyAlaArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 602 CAGGCGGGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661

QY 201 ProLeuVal 203
| | | | |
XX 662 CCCCTGGTG 670
Db
RESULT 14
ABL65450
ID ABL65450 standard; DNA; 1157 BP.
XX
AC ABL65450;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:3787.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 22-SEP-2000; 2000US-234053P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX
PA (AVAL-) AVALON PHARM.
XX

PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX
PS Claim 1; SEQ ID 3787; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1,07e-86 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: Gaps: 0
US-09-827-854-15_COPY_1_203 (1-203) x ABL65450 (1-1157)
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db ATGAAGGCTTCGTGGGCGGTGGCTGCTGTCACATTCCTCGACAGATCCAGGCCAAGGTG 121
QY 21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db GAGCAAGCGGTGCAGACAGACGCCGAGCTGCGCCAGACAGACCGATGGCAGACG 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db GCCAGCGCTGGGAATGCGACATGCGTCTTTGGATTACTGCGCTGGTGACAGACA 241
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db CTGTCTGAGCAGAGTGCAGAGAGCTGCTCAGTCCCAATGCCAAGAACGAGGGCG 301
QY 242 CTGTCTGAGCAGAGTGCAGAGAGCTGCTCAGTCCCAATGCCAAGAACGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnLeu 100
Db CTGTCTGAGCAGAGTGCAGAGAGCTGCTCAGTCCCAATGCCAAGAACGAGGGCG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db ACCCGGTAGCGAGGAGAGCGGGGCGCGCTGCCAAGAGAGTGCAGAGCGGCGAGGCC 421
QY 362 ACCCGGTAGCGAGGAGAGCGGGGCGCGCTGCCAAGAGAGTGCAGAGCGGCGAGGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db CGGCTGGGCGCGAGACATGAGGAGACGTGTGGCGCGCTGTGCAGTACCCGCGGAGGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db CAGGCCATGCTCGGCCAGACACCGAGAGAGTCTGGGTGGCTCGCTCCGCTCCACCTGGCG 541

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 22.2431 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203

Perfect score: 1014

Sequence: 1 MKVLMALVFLAGCOAKV.....AREGAEGLSAIRRLGPLY 203

Scoring table:

BLASTSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/funalt_11032003_101611_27510/app_query.fasta_1.3576
-DB=Issued_Patents.NA -QPMT=fastapi -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdl
-LIST=45 -DOCCALIGN=ptc -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MIMLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn_1.1_22a_etunalt_11032003_101611_27510 -NCPV=6 -ICPV=3
-NO_XLPHY -NO_MMAP -LARGESQUEY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

Issued_Patents.NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCtus.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1005	99.1	1157	1	US-07-709-949-1
2	984	97.0	660	2	US-08-726-306A-28
3	724.5	71.4	1126	4	US-08-949-155-5
4	724.5	71.4	1126	4	US-09-819-964-5
5	588.5	58.0	4267	4	US-08-949-155-51
6	588.5	58.0	4267	4	US-09-819-964-51
7	511	50.4	330	1	US-07-849-389-6
8	374	36.9	252	3	US-08-617-256-24
9	374	36.9	252	4	US-09-287-141-24
10	374	36.9	252	4	US-09-431-613-24
11	374	36.9	252	4	US-09-504-245-24
12	374	36.9	252	4	US-09-287-682-24

13	374	36.9	252	4	US-09-287-679-24	Sequence 24, Appl
14	374	36.9	252	4	US-09-397-766-24	Sequence 24, Appl
15	374	36.9	252	4	US-09-287-681-24	Sequence 24, Appl
16	374	36.9	252	4	US-09-495-444-24	Sequence 24, Appl
17	132	13.0	801	1	US-07-959-946-4	Sequence 4, Appl1
18	132	13.0	801	1	US-08-333-577-4	Sequence 4, Appl1
19	132	13.0	801	5	PCT-US92-08634-4	Sequence 4, Appl1
20	132	13.0	842	4	US-08-952-796-1	Sequence 1, Appl1
21	123	12.1	516	4	US-09-183-861-34	Sequence 34, Appl
22	123	12.1	516	4	US-09-022-765-34	Sequence 34, Appl
23	123	12.1	11236	1	US-07-853-913-1	Sequence 1, Appl1
24	119	11.7	5661	4	US-08-938-105-2	Sequence 1, Appl1
25	118	11.6	592	4	US-09-276-531-115	Sequence 115, App
26	118	11.6	603	4	US-08-056-200-93	Sequence 14, Appl
27	118	11.6	1879	4	US-09-750-580-14	Sequence 2, Appl1
28	116	11.4	863	1	US-08-448-606-7	Sequence 7, Appl1
29	116	11.4	964	1	US-08-448-606-5	Sequence 5, Appl1
30	116	11.4	4296	4	US-09-060-410-3	Sequence 3, Appl1
31	112.5	11.1	3256	2	US-08-968-751-3	Sequence 3, Appl1
32	112	11.0	9551	1	US-08-056-200-93	Sequence 93, Appl
33	112	11.0	9551	2	US-08-800-644-93	Sequence 93, Appl
34	106	10.5	8789	1	US-08-328-254-5	Sequence 5, Appl1
35	106	10.5	10136	1	US-08-353-700-2	Sequence 2, Appl1
36	105.5	10.4	5381	4	PCT-US95-16216-2	Sequence 4, Appl1
37	105.5	10.4	5381	4	US-09-750-580-4	Sequence 1, Appl1
38	105.5	10.4	81001	4	US-08-139-937-12	Sequence 12, Appl
39	105	10.4	4868	1	PCT-US93-11310-12	Sequence 12, Appl
40	105	10.4	4868	5	US-08-195-487-3	Sequence 3, Appl1
41	103.5	10.2	6306	1	PCT-US93-06160-3	Sequence 3, Appl1
42	103.5	10.2	6306	5	US-08-095-737-3	Sequence 3, Appl1
43	102.5	10.1	3033	1	US-08-480-145-3	Sequence 3, Appl1
44	102.5	10.1	3033	2	US-08-477-388-3	Sequence 3, Appl1
45	102.5	10.1	3033	2	US-08-477-388-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5472858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709, 949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

|||||
Db 182 GGCAGCGCTGGAACTGGCACTGGGCTTGGATTAACCTGGCTGGTGACAGACA 241
Qy 61 LeuSerGluGlnValGlnGluGluLeuSerSerGlnValThrGlnGluLeuAla 80
Db 242 CTGTCTAGCAGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
Db 302 CTGATGACACAGACATGAAAGAGTTGAAGCCTACAAATCGGACCTGGAGAACAACTG 361
Qy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
Db 362 ACCCGGTACGAGAGAGACGCGGACGCTCTCCAAAGAGCTGCAGAGCGGCGAGGCC 421
Qy 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGluVal 140
Db 422 CGGCTGGGCGGACATGGAGGAGCTGTGGCGGCTGTGTGACGTAACCGCGGAGGTG 481
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGATGCTGGCCAGACACGAGAGCTCGGGTGCGCTCGCTCCACCTGGCG 541
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
Db 542 AAGCTGGTAAAGGGCTCTCCGCGATCCCGATCCGACCTGCAGAGGCGCTGGCAGTGTAC 601
Qy 181 GlnAlaGlnAlaArgGluGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeu 199
Db 602 CAGGCGGGGCGCGAGGCGCGGCGGCTCAGCGCCATCCGCGAGCGCTG 658

RESULT 3
US-08-949-155-5
; Sequence 5, Application US/08949155
; Patent No. 6271436
; GENERAL INFORMATION:
; APPLICANT: Piedrahita, Jorge A
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/949,155
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 800
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,338
; FILING DATE: 11-OCT-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1126 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 51..1001
US-08-949-155-5
Alignment Scores:
Pred. No.: 1,9e-73 Length: 1126
Score: 724.50 Matches: 146
Percent Similarity: 82.52% Conservative: 24
Best Local Similarity: 70.87% Mismatches: 29
Query Match: 71.45% Indels: 7
Gaps: 2
US-09-827-854-15_COPY_1_203 (1-203) x US-08-949-155-5 (1-1126)
Qy 1 MetLysValLeuThrPheAlaIleLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 51 ATGAGGTTCTGTGGGTTGGCTTTGGTGTACCTCTCCAGAGATCCGACAGAGAC 110
Qy 21 GlnGlnAlaValGluThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 111 GAGCCGGGG-----CCGCGCGCGAGGTGCACGTGTGTGGAGAGGCCAAG 158
Qy 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaGlnGlyArgPheTrpAspTyrLeuArgTrp 57
Db 159 TGGCAGGGCAGCGCCCTGGGAGCAGGCGCTGGCGCTTGGGATTCGTGCGCTGG 218
Qy 58 ValGlnThrLeuSerGlnGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGlu 77
Db 219 GTTCAGTCTCTGTCTGTACCAAGTCAGAGAGAGCTGTCTGACACCAAGTCACCCAGAA 278
Qy 78 LeuArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlu 97
Db 279 CTGACGAGCTGTATAGAGAGACATGAAGAGGTGAAGGCTTACCGCGAGAGCTGGAG 338
Qy 98 GlnGlnLeuThrProValAlaGlnGluThrArgAlaArgLeuSerLysGluLeuGlnAla 117
Db 339 GCGCAGCTGGCGCCCGTACCCAGACGAGAGCGCGGCTGTCCAAAGAGCTGCAGGCG 398
Qy 118 AlaGlnAlaArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArg 137
Db 399 GCGCAGGCGCGCGTGGCGCGCGACATGAGAGAGCTGCGCAACCGCTTGTGTCTACCC 458
Qy 138 GlyGluValGlnAlaMetLeuGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
Db 459 AGCGAGGTGCACAACATGTTGGGCGCAGACCAAGAGAGCTGCGGAGCCGCTGGCTTC 518
Qy 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
Db 519 CACCTCGCAGAGCTGGCGCAGGCGCTGTCGCGACACCGCAGGAGCTGCAGAAAGCGCTG 578
Qy 178 AlaValTyrGlnAlaGlnAlaArgGluGlnAlaGluArgGlyLeuSerAlaIleArgGlu 197
Db 579 GCGGTGTACAGGCGGGGCTGGCGGAGGCGCGGAGCGCGAGCGTGTGAGCGCTCTCGGAG 638
Qy 198 ArgLeuGlyProLeuVal 203
Db 639 CGCCTCGGCGCCCTGGTG 656

RESULT 4
US-09-819-964-5
; Sequence 5, Application US/0981964
; Patent No. 6369294
; GENERAL INFORMATION:
; APPLICANT: Piedrahita, Jorge A
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Generation of Transgenic Animal Species
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/819,964
FILING DATE: 28-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,155
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-May-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMK-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
SEQUENCE DESCRIPTION: SEQ ID NO: 5
US-09-819-964-5
Alignment Scores:
Pred. No.: 1,9e-73 Length: 1126
Score: 724.50 Matches: 146
Percent Similarity: 82.52% Conservative: 24
Best Local Similarity: 70.87% Mismatches: 29
Query Match: 71.45% Indels: 7
Gaps: 2
US-09-827-854-15_COPY_1_203 (1-203) x US-09-819-964-5 (1-1126)
QY 1 MettysValleutTPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 51 ATGAGGGTCTGTGGTGTCTTGGTGTAACTCTCGAGAGATCCGAGACAGAGAC 110
QY 21 GtUGlnAlaValGtUthrGluProGluProGluLeuArg-----GlnGlnThGlu 37
Db 111 GAGCGGGG-----CCGCCCGAGAGGTGCACGTGTGGTGGAGAGAGCCCAAG 158
QY 38 TrpGlnserGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
Db 159 TGGCAGGGCAGCAGCCCTGGGAGACAGCCCTGGCGCTTCTGGATTACCTCGCTGG 218
QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
Db 219 GTGCAGTCCCTGTGTACCAAGTGCAGAGAGCTGCTCAGACACAGAGTCCACCCAGAA 278
QY 78 LeuArgAlaLeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlu 97
Db 279 CTGACGAGCTGATAGAGGAGCATGAGAGGTGAAGGCTTACCCGAGAGAGCTGGAG 338
QY 98 GlUGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerLysGlnLeuGlnAla 117
Db 339 GCGCAGCTGGCCCGCTGACCCAGAGACAGCAGCGCCCTGTCTCAAGAGAGCTGACGGC 398

QY 118 AlaGlnAlaArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArg 137
Db 399 GCGCAGGCCCGCGGTGGCGGACATGAGAGACGTGGCAACCCCTTGCTCTTACCGC 458
QY 138 GlyGluValGlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSer 157
Db 459 AGCAGGTGCACACATGTTGGGCGACACACAGAGAGCTGGCAGCCGCTTGCTCC 518
QY 158 HisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 177
Db 519 CACCTGGCAAGCTGCGCAAGCGGCTGCTCGCGACACCGAGCAGCAGCAAGACGCGCTG 578
QY 178 AlaValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlu 197
Db 579 GCCCTGTACCAAGCGGGCGTGGCGAGGCGCCGAGCGCAGCAGCGTGCCTCCGCGAG 638
QY 198 ArgLeuGlyProLeuVal 203
Db 639 GCGCTCGGGCCCTGTGTG 656
RESULT 5
US-08-949-155-51
Sequence 51, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-Oct-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-May-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMK-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 4267 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-949-155-51
Alignment Scores:
Pred. No.: 3.82e-57 Length: 4267
Score: 588.50 Matches: 140
Percent Similarity: 48.94% Conservative: 22
Best Local Similarity: 42.30% Mismatches: 31
Query Match: 58.04% Indels: 139

DB: 4 Gaps: 4
US-09-827-854-15_COPY_1_203 (1-203) x US-08-949-155-51 (1-4267)
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 2448 ATGGCTGTAACTGGCTGCGT-----GCAGATGCGGACAGAGGAC 2489
QY 21 GluGlnAlaValGluThrGluProGluLeuArg-----GlnGlnThrGlu 37
DB 2490 GAGCGGGG-----CCGCGCGGAGGTGCACGTGTGGTGGAGGAGCCCAAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
DB 2538 TGGCAGGGGACGACCCCTGGGAGAGGCCCTTGTGGGATTACTGCGCTGG 2597
QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 2598 GTGCAGTCCCTGTCTGACCAAGTGCAGAGAGAGCTGCACCAAGTCCACCGAGA 2657
QY 78 Leu----- 78
DB 2658 CT-GACGTAAAGTCCACCCGACTCCCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
DB 2717 CCCTCTGGGCAACCGTGTCTGTGACCCCTCAGCCTCCGCGGTTCTCTGTG 2776
QY 78 ----- 78
DB 2777 TCCCTGTGCGCAACTCTTGGGGGTCTGGGTCTGTCTTCTTCTTCTTCTTCTT 2836
QY 78 ----- 78
DB 2837 TTGGGGGGAATACTTTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCC 2896
QY 78 ----- 78
DB 2897 ATCTTGTAGCTCTGCTTGCCTGTCTGTGGTCTGAGTCTTGGCGCTTGTCTGTG 2956
QY 78 ----- 78
DB 2957 AATCTGTGGCAGTCTGTGGCCATGCCAGCTCAGAGAGCCCTTCTCCCTCACCAGC 3016
QY 79 ----- 79
DB 3017 CCGCGCTCTCTGCGCCCAAGGAGCTGATAGAGAGATGAGAGAGGTAGAGCCTAC 3076
QY 93 LysSerGluLeuGluGlnGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSer 112
DB 3077 CCGCAGAGAGCTGGAGCGAGCTGGCCCGCGTACCCAGAGAGAGAGCGCGCTGTCC 3136
QY 113 LysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAspValCysGlyArg 132
DB 3137 AAGGAGCTGACAGCGCGCCAGCGCGCTGGCGCGCCGATGTGAGAGAGTGGCAACCGC 3196
QY 133 LeuValGlnTyrArgGlyValGlnAlaMetLeuGlyGlnSerThrGlnGlnLeuArg 152
DB 3197 TTGGTGTCTTACCCAGCAGGAGGTGCACAACATGTTGGCCAGACACCGAGAGAGTGGCG 3256
QY 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp 172
DB 3257 AGCGCGCTGTCTCCACCTGTGGCAAGCTGGCGCAAGCGCTGTCTCCGAGACACGAGAG 3316
QY 173 LeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAlaGluArgLysLeu 192
DB 3317 CTGCAGAAAGCGCTGGCGCTGTACACAGCGGGGCTGCGCGAGGGCGCGAGCGCAGCGTG 3376
QY 193 SerAlaIleArgGluArgLeuGlyProLeuVal 203
DB 3377 AGCGCCTTCCGCGAGCGCTCGGGCCCTGTGGT 3409

RESULT 6

US-09-819-964-51
; Sequence 51, Application US/09819964
; Patent No. 6369294
; GENERAL INFORMATION:
; APPLICANT: Piedrahita, Jorge A
; TITLE OF INVENTION: Bazer, Fuller W
; Generation of Transgenic Animal Species
; NUMBER OF SEQUENCE ADDRESSES: 51
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: ARNOLD, WHITE AND DURKEE
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/819,964
; FILING DATE: 28-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/949,155
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 60/046,094
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: TANK:177
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4267 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-819-964-51
Alignment Scores:
Pred. No.: 3,82e-57 Length: 4267
Score: 588.50 Matches: 140
Percent Similarity: 48.94% Conservative: 22
Best Local Similarity: 42.30% Mismatches: 31
Query Match: 58.04% Indels: 139
DB: 4 Gaps: 4
US-09-827-854-15_COPY_1_203 (1-203) x US-09-819-964-51 (1-4267)
QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 2448 ATGGCTGTAACTGGCTGCGT-----GCAGATGCGGACAGAGGAC 2489
QY 21 GluGlnAlaValGluThrGluProGluLeuArg-----GlnGlnThrGlu 37
DB 2490 GAGCGGGG-----CCGCGCGGAGGTGCACGTGTGGTGGAGGAGCCCAAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
DB 2538 TGGCAGGGGACGACCCCTGGGAGAGGCCCTTGTGGGATTACTGCGCTGG 2597
QY 58 ValGlnThrLeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGlu 77
DB 2598 GTGCAGTCCCTGTCTGACCAAGTGCAGAGAGAGCTGCACCAAGTCCACCGAGA 2657
QY 78 Leu----- 78

```
Db 2658 CT-GACGTAGTCCACCCAGCTCCCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCCTCTGGGGAACCGTGTCTTGACACCTCAGAGCTCCACCCGTCGGGTTCTTCTG 2776
QY 78 ----- 78
Db 2777 TCCTGTCCGCACTCTGGGGGTCTGGTCTGTTCTTTCTTTCTTTCTTCTTCTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGAAAAAATTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTCC 2896
QY 78 ----- 78
Db 2897 ATCTTAGCTCTGCTCTGCGCTGTCTGTGGTCTGAGTCTTGCGCTTCTGCTCTG 2956
QY 78 ----- 78
Db 2957 AATCTGTGACGCTGCTGGCATCGCAGCTCAGAGCCCTCTCTGCCCCCTCACCGCC 3016
QY 79 ----- 79
Db 3017 CCGCCCTCTCTGCGCCAGGAGCTGATAGAGAGAGATGAAGAGAGTGAAGGCTAC 3076
QY 93 LysSerGluLeuGluGluGluLeuThrProValAlaGluGluThrArgLeuSer 112
Db 3077 GCGGAGAGCTGGAGGCGCGCTGGGCGCCGTGACCCAGAGAGAGAGGCGCCCTGTTC 3136
QY 113 LysGluLeuGluAlaAlaGluAlaArgLeuGlyAlaAspMetGluAspValCysGlyArg 132
Db 3137 AAGGAGCTGACGAGCGGCGGAGCGCCGCTGGGGCGGACATGAGAGACGTCGCAACCGC 3196
QY 133 LeuValGluIleThrArgGlyGluValGluAlaMetLeuGlyGlnSerThrGluGluLeuArg 152
Db 3197 TTGGTCTCTACCGCAGCGAGGTGCACAACTGTGGCGCAGCACCGGAGGCTGCGG 3256
QY 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAsp 172
Db 3257 AGCGGCTGCTTCCCACTCGCGCAAGCTGCGCAAGCGGCTGCTCCGCGCACCGGAGAC 3316
QY 173 LeuGluLysArgLeuAlaValTyrglnAlaGlyAlaArgGluGlyAlaGluArgGlyLeu 192
Db 3317 CTCAGAAAGCGCTGCGCTGTACAGCGGGGCTCGCGAGGGCGCGCAGCGCACGCTG 3376
QY 193 SerAlaIleArgGluArgLeuGlyProLeuVal 203
Db 3377 AGCGCCCTCGCGAGCGCTCGGGCCCTGTGTG 3409

RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5325493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22133-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/07/849,389
FILING DATE: 19920519
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 330 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: HUMAN LIPOPROTEIN E GENE
FEATURE:
NAME/KEY: CDS
LOCATION: 1..330
US-07-849-389-6

Alignment Scores:
Pred. No.: 6,84e-50 Length: 330
Score: 511.00 Matches: 107
Percent Similarity: 98.18% Conservative: 1
Best Local Similarity: 97.27% Mismatches: 2
Query Match: 50.39% Indels: 0
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-07-849-389-6 (1-330)

QY 87 LysGluLeuLysAlaTrpLysSerGluLeuGluGluGluLeuThrProValAlaGluGlu 106
Db 1 AAGGAGTTGAAGCCCTACAAATCGGAATCGAGGAACAACCTACCCCGTGGCGGAGAG 60
QY 107 ThrArgAlaArgLeuSerLysGluLeuGluGluAlaAlaGluAlaArgLeuGlyAlaAspMet 126
Db 61 AGCGGCGCAGCGCTGTCCAAAGAGCTGCAGCGCGGAGGAGCCCGCTGGGCGGACATG 120
QY 127 GluAspValCysGlyArgLeuValGlnTyrglnArgGlyGluValGluAlaMetLeuGlyGln 146
Db 121 GAGGAGCTGCGCGCGCGCTGTGTGACGTACCGCGGCGGAGGTGCAGGCTGCGGCGAG 180
QY 147 SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeu 166
Db 181 AGCAGCGAGAGCTGCGGGGTGCGCTCGCTCCACCTCGCAGAGCTGCTGAAGCGGCTC 240
QY 167 LeuArgAspAlaAspAspLeuGluLysArgLeuAlaValTyrglnAlaGluAlaArgGlu 186
Db 241 CTCGCGGATCGCGATGACTGCAGAAAGCCCTGCGCATGTACAGGCGCGGCGCGGAG 300
QY 187 GlyAlaGluArgGlyLeuSerAlaIleArg 196
Db 301 GGGCGCGAGCGGCGCTCAGCGCATCCGC 330

RESULT 8
US-08-617-256-24
; Sequence 24, Application US/08617256
; Patent No. 6043031
; GENERAL INFORMATION:
; APPLICANT: Kyster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
```

ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,256
FILING DATE: March 18, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: March 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth A.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: SOL-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-617-256-24

Alignment Scores:
Pred. No.: 1,86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 3
Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-08-617-256-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGTCGACGGCGCGCGCGCTGGCGGCACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCG-CGCTGTGTGAGTACCGCGCGAGGTCGACGCGCATGCTCGGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCCCTCCACCTGCGCAAGCTGCGTAAAGCGGCTCCCTCCGC 180
QY 169 AspAlaAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGCAAGTACCAAGCGCGGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCTC 252
RESULT 9
US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCaulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA

ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-287-141-24

Alignment Scores:
Pred. No.: 1,86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 4
Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-141-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGTCGACGGCGCGCGCGCTGGCGGCACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCG-CGCTGTGTGAGTACCGCGCGAGGTCGACGCGCATGCTCGGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCTCCCTCCACCTGCGCAAGCTGCGTAAAGCGGCTCCCTCCGC 180
QY 169 AspAlaAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGCAAGTACCAAGCGCGGCGCGCGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGGGCTC 252
RESULT 10
US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-450-8400
TELEFAX: 619-587-5360
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-431-613-24 (1-252)
QY 109 AAlArGleUSeRlYsglUleuGlnAlaGlnAlaArGleuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGGCGCCAGCCCGCTGGGGCGGAGCATGGAGAC 61
QY 129 ValCysglYArGleuValGlnTYrArGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-GGCTGTGGTGCAGTACCGGGCGAGGTGCAGGCATGTCTCGCCGAGACACC 120
QY 149 GluGlUleuArGyValArGleuAlaSerHisLeuArGlyLeuArGlySArGleuLeuArg 168
DB 121 GAGGAGCTGGGGGTGGCCCTCGCCCTCCACCTGGCGCAAGCTGGCTAAGCGGCTCTCCG 180
QY 169 AspAlaAspAspLeuGlnTySArGleuAlaValTYrGlnAlaGlyAlaArGlyGluGlyAla 188
DB 181 GATGCCGATACCTGCAGAGAGTCCCTGTGCAGTGTACACAGCGGGGCGCCGAGAGCGCC 240
QY 189 GluArGlyLeu 192
DB 241 GAGCGCGGCTC 252
RESULT 11
US-09-504-245-24
```

```

Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McCauliffe LLP
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/504,245
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858-450-8400
TELEFAX: 858-587-5360
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-504-245-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-504-245-24 (1-252)
QY 109 AAlArGleUSeRlYsglUleuGlnAlaGlnAlaArGleuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAAAGAGCTGCAGGGCGCCAGCCCGCTGGGGCGGAGCATGGAGAC 61
QY 129 ValCysglYArGleuValGlnTYrArGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTGGCC-GGCTGTGGTGCAGTACCGGGCGAGGTGCAGGCATGTCTCGCCGAGACACC 120
QY 149 GluGlUleuArGyValArGleuAlaSerHisLeuArGlyLeuArGlySArGleuLeuArg 168
DB 121 GAGGAGCTGGGGGTGGCCCTCGCCCTCCACCTGGCGCAAGCTGGCTAAGCGGCTCTCCG 180
QY 169 AspAlaAspAspLeuGlnTySArGleuAlaValTYrGlnAlaGlyAlaArGlyGluGlyAla 188
DB 181 GATGCCGATACCTGCAGAGAGTCCCTGTGCAGTGTACACAGCGGGGCGCCGAGAGCGCC 240
QY 189 GluArGlyLeu 192
DB 241 GAGCGCGGCTC 252
```

RESULT 12
US-09-287-682-24
; Sequence 24, Application US/09287682
; Patent No. 6235478
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; APPLICANT: Little, Daniel P.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & Mcauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-287-682-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-682-24 (1-252)
QY 109 AAlaArgLeuSerIySGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGCGCCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyInserThr 148
DB 62 GTGTGGCC-CGGCTGTGTGAGTACCGCGCGGCGAGGTCCAGCCATGCTGGCCAGAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgGlyLeuArgGlySerArgLeuArg 168
DB 121 GAGAGACTCGGGGTGGCTCGCTCCACCTGCGCAGAGCTCGTAAAGCGGCTCTCCGCC 180

QY 169 AspAlaAspLeuGlnIlySerArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACTGCAGAGATCCCTGCGCATGTACCAAGCGCGGCCCGCGAGGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 13
US-09-287-679-24
; Sequence 24, Application US/09287679
; Patent No. 6258538
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; APPLICANT: Little, Daniel P.
; APPLICANT: Braun, Andreas
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & Mcauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,679
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-09-287-679-24

Alignment Scores:
Pred. No.: 1,86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-287-679-24 (1-252)
QY 109 AAlaArgLeuSerIySGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluasp 128
DB 2 GCACGGCTGTCCAGAGAGCTGCAGGGCGCGCCGCTGGCGCGGACATGGAGGAC 61

QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
| | | | |
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCATGTCTCGGACAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
| | | | |
Db 121 GAGAGCTGGGGTGGCCCTGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspaAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
| | | | |
Db 181 GATGCCGATACCTGCGAAGTCCCTGTGCAAGTGTACAGGCGCGGCGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
| | | | |
Db 241 GAGCGCGGCTC 252
RESULT 14
US-09-397-766-24
; Sequence 24, Application US/09397766
; Patent No. 6268144
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/397,766
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-20021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-397-766-24
Alignment Scores:
Pred. No.: 1.86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-397-766-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
| | | | |
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCGGCGGCGGCGGAGCATGTGAGAGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
| | | | |
Db 62 GTGTGGCC-CGCTGTGTGAGTACCGCGGAGGTGCAGGCATGTCTCGGACAGAGCACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
| | | | |
Db 121 GAGAGCTGGGGTGGCCCTGCTCCACCTGCGCAAGCTGCTAAGCGGCTCTCCGC 180
QY 169 AspaAlaAspAspLeuGlnTyrArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
| | | | |
Db 181 GATGCCGATACCTGCGAAGTCCCTGTGCAAGTGTACAGGCGCGGCGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
| | | | |
Db 241 GAGCGCGGCTC 252
RESULT 15
US-09-287-681-24
; Sequence 24, Application US/09287681
; Patent No. 6277573
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,681
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-287-681-24
Alignment Scores:
Pred. No.: 1.86e-34 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 36.88% Indels: 1
DB: 4 Gaps: 0

Pred. No.: 1,86e-34
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 36.88%
DB: 4

Length: 252
Matches: 82
Conservative: 0
Mismatches: 2
Indels: 1
Gaps: 0

US-09-827-854-15_copy_1_203 (1-203) x US-09-287-681-24 (1-252)

QY 109 AlaArgLeuSerLysGluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAGAGCTGCAGCGCGCGCCGCTGGCGCGGACATGGAGGAC 61
QY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGCGC-CGCTGTGTGTCAGTACCGCGCGAGGTGCAGGCCATGTCTGGCCAGACACC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuLeuArg 168
Db 121 GAGGAGCTGGGGTGCCTCGCTCCACCTGCGCAAGCTGCTAAGCGGCTCCTCCGCC 180
QY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGAGACCTGCAGAAATCCCTGGCAGTGTACAGGCCGGGGCCCGCGAGGGCGGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCTC 252

Search completed: March 14, 2003, 20:18:35
Job time : 26.2431 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 42.1851 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203
Perfect score: 1014
Sequence: 1 MKVLMALVTLFLGCAKAY.....ARCAERGLSAIRERLGLV 203

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPNO.spool/US09827854/rnat_11032003_101612_27578/app_query.fasta.1.3576
-DB=Published.Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09827854@CGN.1.1.298 -runat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/USN7_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PC2_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCRTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1014	100.0	1156	9	US-09-870-759-129
2	1014	100.0	1156	9	US-09-802-640-17
3	1014	100.0	1156	10	US-09-827-854-8
4	1014	100.0	1291	12	US-10-044-090-454

5	1010	99.6	1156	10	US-09-827-854-12	Sequence 12, App1
6	1006	99.2	1156	10	US-09-827-854-9	Sequence 9, App1
7	1006	99.2	1156	10	US-09-827-854-11	Sequence 11, App1
8	1005	99.1	1157	10	US-09-954-456-760	Sequence 760, App
9	1005	99.1	1157	10	US-09-880-107-2244	Sequence 2244, App
10	1002	98.8	1156	10	US-09-827-854-7	Sequence 7, App1
11	999	98.5	1156	10	US-09-827-854-10	Sequence 10, App1
12	798.5	78.7	41907	10	US-09-967-013-5	Sequence 5, App1
13	625	61.6	478	10	US-09-964-824A-374	Sequence 374, App
14	625	61.6	478	10	US-09-880-107-2491	Sequence 2491, App
15	458.5	45.2	449	10	US-09-960-352-4726	Sequence 4726, App
16	458	45.2	356	10	US-09-960-352-5420	Sequence 5420, App
17	439.5	43.3	442	10	US-09-960-352-9395	Sequence 9395, App
18	435	42.9	414	10	US-09-960-352-4237	Sequence 4237, App
19	428.5	42.3	416	10	US-09-960-352-8720	Sequence 8720, App
20	422.5	41.7	423	10	US-09-960-352-14047	Sequence 14047, App
21	405.5	40.0	409	10	US-09-960-352-5148	Sequence 5148, App
22	401.5	39.6	416	10	US-09-960-352-9797	Sequence 9797, App
23	401	39.5	253	10	US-09-179-536B-130	Sequence 130, App
24	387.5	38.2	425	10	US-09-960-352-3497	Sequence 3497, App
25	380.5	37.5	390	10	US-09-960-352-1311	Sequence 1311, App
26	378.5	37.4	391	10	US-09-960-352-1278	Sequence 1278, App
27	379.5	37.4	392	10	US-09-960-352-10599	Sequence 10599, App
28	379.5	37.4	401	10	US-09-960-352-4187	Sequence 4187, App
29	375.5	37.0	413	10	US-09-960-352-8042	Sequence 8042, App
30	374	36.9	252	10	US-09-960-352-8416-24	Sequence 24, App1
31	374	36.9	252	10	US-09-879-341-24	Sequence 24, App1
32	369.5	36.4	404	10	US-09-960-352-2325	Sequence 2325, App
33	369.5	36.4	405	10	US-09-960-352-14063	Sequence 14063, App
34	351.5	34.7	377	10	US-09-960-352-176	Sequence 176, App
35	350.5	34.6	353	10	US-09-960-352-4914	Sequence 4914, App
36	350.5	34.6	378	10	US-09-960-352-10040	Sequence 10040, App
37	350.5	34.6	386	10	US-09-960-352-10361	Sequence 10361, App
38	348.5	34.4	388	10	US-09-960-352-11986	Sequence 11986, App
39	348.5	34.1	396	10	US-09-960-352-3457	Sequence 3457, App
40	344.5	34.0	377	10	US-09-960-352-10540	Sequence 10540, App
41	338.5	33.4	432	10	US-09-960-352-8934	Sequence 8934, App
42	338	33.3	372	10	US-09-960-352-6552	Sequence 6552, App
43	337.5	33.3	376	10	US-09-960-352-5408	Sequence 5408, App
44	335.5	33.1	373	10	US-09-960-352-3681	Sequence 3681, App
45	334.5	33.0	369	10	US-09-960-352-3494	Sequence 3494, App

ALIGNMENTS

RESULT 1
US-09-870-759-129
Sequence 129, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870, 759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 129
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION:
US-09-870-759-129

Alignment Scores: 3.37e-100 Length: 1156
Pred. No.: 1014.00 Matches: 203
Score: 1014.00
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

Mismatches: 0
Indels: 0
Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-870-759-129 (1-1156)

```
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||
DB 61 ATGAAGTTCTGTGGCTGCTGTGCTGCATCTCTCGCAGAGATCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
    |||||
DB 121 GACCAAGCGGTGGAGACAGAGCCGAGCCGACCTGGCCAGACAGACCGAGTGGCAGAGC 180
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||
DB 181 GGCACAGCGTGGAGACTGGCAGCTGGCTTTTGGGATTAACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||
DB 241 CTGTCTGAGCAGGTGACAGAGAGCTGTCTCAGCTCCAGGTACCCAGAACTGAGGGCC 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
    |||||
DB 301 CTGATGGACGAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||||
DB 361 ACCCGGTTGGCGGAGAGAGCGCGGCTGTCCAGAGAGCTGCAGGGCGGCGCAGGCC 420
QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGluVal 140
    |||||
DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGGCTGTGACAGTACCAGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
    |||||
DB 481 CAGGCCATGCTCGCCACAGACACCGAGAGCTCGGGTGGCTCCCTCCCTCCACTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
    |||||
DB 541 AACCTGCGTAAAGGCGCTCTCCCGCATGCCGATGACCTGACAGAAAGCCCTGGCAGTTC 600
QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||
DB 601 CAGGCCGGGGCGCGCAGAGCGCGGCGGCTCAGCCATCCGCGAGCGCCCTGGGG 660
QY 201 ProLeuVal 203
    |||||
DB 661 CCCCTGGTG 669
```

RESULT 2
US-09-802-640-17
Sequence 17, Application US/09802640
Publication No. US20030036057A1

GENERAL INFORMATION:

APPLICANT: Braun, Andreas
APPLICANT: Bonsal Aruna
APPLICANT: Kleya Patrick
TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: 24736-2048
CURRENT APPLICATION NUMBER: US/09/802,640
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E

OTHER INFORMATION: (APOE)
US-09-802-640-17

Alignment Scores:

Pred. No.:	3 37e-100	Length:	1156
Score:	1014.00	Matches:	203
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-802-640-17 (1-1156)

```
QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
    |||||
DB 61 ATGAAGTTCTGTGGCTGCTGTGCTGCATCTCTCGCAGAGATCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnInThrGluTrpGlnSer 40
    |||||
DB 121 GACCAAGCGGTGGAGACAGAGCCGAGCCGACCTGGCCAGACAGACCGAGTGGCAGAGC 180
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
    |||||
DB 181 GGCACAGCGTGGAGACTGGCAGCTGGCTTTTGGGATTAACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||
DB 241 CTGTCTGAGCAGGTGACAGAGAGCTGTCTCAGCTCCAGGTACCCAGAACTGAGGGCC 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
    |||||
DB 301 CTGATGGACGAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACCTGGAGAACACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
    |||||
DB 361 ACCCGGTTGGCGGAGAGAGCGCGGCTGTCCAGAGAGCTGCAGGGCGGCGCAGGCC 420
QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGluVal 140
    |||||
DB 421 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGGCTGTGACAGTACCAGCGGAGAGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValAlaArgLeuAlaSerHisLeuArg 160
    |||||
DB 481 CAGGCCATGCTCGCCACAGACACCGAGAGCTCGGGTGGCTCCCTCCCTCCACTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
    |||||
DB 541 AACCTGCGTAAAGGCGCTCTCCCGCATGCCGATGACCTGACAGAAAGCCCTGGCAGTTC 600
QY 181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
    |||||
DB 601 CAGGCCGGGGCGCGCAGAGCGCGGCGGCTCAGCCATCCGCGAGCGCCCTGGGG 660
QY 201 ProLeuVal 203
    |||||
DB 661 CCCCTGGTG 669
```

RESULT 3
US-09-827-854-18
Sequence 8, Application US/09827854
Patent No. US20020123093A1

GENERAL INFORMATION:

APPLICANT: Zannis, Vassilis
APPLICANT: Kyriacos, Kyriakos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
PRIORITY FILING DATE: 2001-04-05
PRIORITY FILING DATE: US 09/679,088
PRIORITY FILING DATE: 2000-10-04
PRIORITY FILING DATE: US 09/544,386
PRIORITY FILING DATE: 2000-04-06

QY 201 ProLeuVal 203
Db 661 CCCCTGGTG 669

RESULT 7
US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Zannis, Vassilis
; APPLICANT: Kyriacos, Kyriacos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; TITLE OF INVENTION: Cholesterol levels without inducing hypertriglyceridemia
; FILE REFERENCE: 07180/004003
; CURRENT APPLICATION NUMBER: US/09/827,854
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-11

Alignment Scores:
Pred. No.: 2,44e-99 Length: 1156
Score: 1006.00 Matches: 202
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 99.21% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-827-854-11 (1-1156)

QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
Db 61 ATGAAGGTTCTGTGGCTGCTGGTGGTGCACATTCCTGGCGAGTGCAGGCAAGGTG 120

QY 21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCCGCGCCAGACAGCCAGGCGCAGACC 180

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 181 GGCACGCGCTGGAACTGCTGCTGCTTTGGATTACCTGCGCTGGGTGCAGACA 240

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGCTTGACGAGGTGAGAGAGGCTGCTCAGCTCCAGGTCCAGGAGTGAAGGCGG 300

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100
Db 301 CTGATGAGACGAGACCATGAGAGGCTTGAAGGCTTCAATCGACTGGAGGAACTG 360

QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
Db 361 ACCCGCGGTGGGAGAGACGCGGCGCTGCTCAAGAGGCTCAGGCGCGCAGGCC 420

QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
Db 421 CGGCTGGGCGGAGATGAGAGACGTGCGCGCGCTGGGTGCACTACCGGCGCAGGTG 480

QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCGCATGCTGCGCAGAGACGAGAGCTGGGGGTGCGCGCTCCACCACTGGCGC 540

QY 161 LysLeuAlaGlyLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGTAAAGCGGCTCTCCGCGATGCGATGACCTGAGAGAAAGCGCGCTGCGAGTTC 600

QY 181 GlnAlaGlyAlaArgGluValArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGGCCGGGGCCCGGAGGGCCGCGAGCCGGCTCAGCGCATCCGCGAGCGCTGGGG 660

QY 201 ProLeuVal 203
Db 661 CCCCTGGTG 669

RESULT 8
US-09-954-456-760
; Sequence 760, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 760
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-760

Alignment Scores:
Pred. No.: 3.13e-99 Length: 1157
Score: 1005.00 Matches: 201
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 2
Query Match: 99.11% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-954-456-760 (1-1157)

QY 1 MetLysValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
Db 62 ATGAAGGTTCTGTGGCTGCTGGTGGTGCACATTCCTGGCGAGTGCAGGCAAGGTG 121

QY 21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 122 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCCAGACAGCGAGTGCAGAGC 181

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 182 GGCACGCGCTTGGAACTGCTGCTGCTTTGGATTACCTGCGCTGGGTGAGACA 241

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerGlnValThrGlnGluLeuArgAla 80
Db 242 CTGCTGAGCAGGTGACAGAGAGCTGCTCAGCTCCCAAGTCCACCAAGAACTGAGGCG 301

Alignment Scores:	
Pred. No.:	3,13e-99
Score:	1005.00
Percent Similarity:	99.01%
Best Local Similarity:	99.01%
Query Match:	99.11%
DB:	10
US-09-827-854-15_COPY_1_203 (1-203) x US-09-880-107-2244 (1-1157)	
QY	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db	62 ATGAAGGTTCTGTGGCGTCGCTTGCTGCATTCCTCGCAGAGATGCCAGCCAGAGTG 121
QY	21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db	122 GAGCAAGCGGTGTGAGACAGAGACCCGAGCCCGAGCTGGCCGACGAGACCGAGTGGCAGAGC 181

```

RESULT 10
US-09-827-854-7
; Sequence 7, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:
; APPLICANT: Kypros, Kyriakos E.
; TITLE OF INVENTION: Compounds and methods for lowering
; FILE REFERENCE: 07180/004003
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: US 09/679,088
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-7

Alignment Scores:
Pred. No.: 6.57e-99 Length: 1156
Score: 1002.00 Matches: 202
Percent Similarity: 99.51% Conservativity: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 98.82% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-827-854-7 (1-1156)
QY 1 MctlysvalettRPaLaLaLaLeuVaLtrPheneuLaLaLlyGyGgnatLaLysval 20
|||||

```


Query Match: 78.758 Indels: 194
DB: 10 Gaps: 1

US-09-827-854-15_COPY_1_203 (1-203) x US-09-967-013-5 (1-41907)

OY 12 PheLeuAlaGlycysGlnAlaLysValGluGlnAlaValGluThrGluProGluProGlu 31
DB 20316 TTCACACAGGATGCCAGCAAGGTGGAGCAAGCGGTGAGACAGACGCCGAGAGCCCGGAG 20375

OY 32 LeuArgGlnGlnThrGluThrGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPhe 51
DB 20376 CTGGGCGACAGACGACGATGGCGAGAGCGGCGACGCTGGGAATCGGCACGCGGCTTT 20435

OY 52 TrpAspTrpLeuArgTrpValGlnThrLeuSerGluGlnValGlnGluLeuLeuSer 71
DB 20436 TGGGATTTACCTGGCGTGGGGGACAGACTGTCTGACAGAGTGCGAGAGAGAGCTCAGC 20495

OY 72 SerGlnValThrGlnGlnLeu----- 78
DB 20496 TCCGAGGTACCCAGAGAACTGAGGTAGTGTCCCATCTGGCCCTTGACCTCTGTGTG 20555

OY 78 ----- 78
DB 20556 GGGGGGTATACCTCCAGGTCCAGGTTTCATCTGCCCCCTGTGCGTAAGTCTGGGGGG 20615

OY 78 ----- 78
DB 20616 CTTGGGCTCTGCTGCTGCTAGCTCTCTCTCCATTTCTGACTCTGCTTACGCTTC 20675

OY 78 ----- 78
DB 20676 TGGAAATCTCTCTCTGACGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20735

OY 78 ----- 78
DB 20736 CGTCTGGCTCTGCT 20795

OY 78 ----- 78
DB 20796 CTGCTGTGTGGCCAGGCTGTGCTTGAACCTCTGGGCTTCAGACGATCTCTCTCTCT 20855

OY 78 ----- 78
DB 20856 CTTCCCAAAAGTCTGGGATTTAGAGCATGAGCACCTTGGCCGCTCTAGCTCTCTCT 20915

OY 78 ----- 78
DB 20916 TCGTCTCTGCT 20975

OY 78 ----- 78
DB 20976 GCTCTGCCCCGTTCT 21035

OY 78 ----- 78
DB 21036 CCCCATCCAGCCCTTCTCCCGCCCTCCACCTGTGCGACACCTCCCGCCCTCTCGCGC 21095

OY 79 -ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGln 98
DB 21096 CAGGCGCGTGTGAGAGACCATGAAGAGTTGAAGGCTTCAATCGGAATCGGAGAG 21155

OY 98 uGlnLeuThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAla 118
DB 21156 ACAACTGACCCCGGTGGCGAGAGACGCGGCGACGCGCTGTCCAAAGACTGACGAGCG 21215

OY 118 aGlnAlaArgLeuGlyAlaAspMetLysPheValLysGlyArgLeuValGlnTrpArg 138
DB 21216 GGAGGCGCGGTGGCGGCGAGATGAGAGCGTGGCGCGCTGTGTGTCAGATCGCGG 21275

OY 138 yGluValGlnAlaMetLeuGlyGlnSerThrGluGlnLeuArgValArgLeuAlaSer 158
DB 21276 CGAGGTGACGCGCATCTCTGCGCCAGAGACCAAGAGAGCTGCGGCTCTCTCTCTCT 21335

OY 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeu 178
DB 21336 CTTGCCCAAGCTGTGTAAAGGCGCTCTCCGCGATGCGGATGACCTCGAGAAAGCGCTGGC 21395

OY 178 aValTrpGlnAlaGlyAlaArgGluGlnGlyAlaArgGlyLeuSerAlaIleArgGln 198
DB 21396 AGTGTACAGAGCGCGGCGCGCGAGGCGCGGAGCGCGGCTCAGCGCCATCGCGAGCG 21455

OY 198 gLeuGlyProLeuVal 203
DB 21456 CTGGGCGCCCTGTGTG 21471

RESULT 13
US-09-964-824A-374/c
; Sequence 374, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT FILING DATE: 2001-09-27
; PRIOR FILING DATE: US/60/236,033
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 374
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(478)
; OTHER INFORMATION: n=a,t,g or c
US-09-964-824A-374

Alignment Scores:
Pred. No.: 6,98e-59 Length: 478
Score: 625.00 Matches: 126
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 96.92% Mismatches: 2
Query Match: 61.64% Indels: 1
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x US-09-964-824A-374 (1-478)

OY 1 MetLysValLeuThrProAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 396 ATCAAGTCTCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 337

OY 21 GlnGlnAlaValGlnThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 336 GAGCAAGCGGTGGAGAGACAGAGCGGAGCCGACGCTGCGCAGAGAGCCAGATGGCGAGAGC 277

OY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 276 GGCACAGCGTGGAACTGGCAGTGGGCTGTGGGATTTACCTGCGCTGGGTGCGAGACA 217

OY 61 LeuSerGlnGlnValGlnGlnGluLeuLeu-SerSerGlnValThrGlnGluLeuArgAl 80
DB 216 CTGTCTGACAGAGTGCAGAGAGAGAGCTGCTGAGCTCCAGAGTCCAGCAGAACTGAGGCG 157

OY 80 aLeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGln 100
DB 156 GCTGATGAGCAGAGACCATGAAAGAGTTGAAGGCTTCAATCGGAATCGGAGAACAACT 97

OY 100 uThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120
DB 100 uThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAl 120

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 825.891 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_203

Perfect score: 1014

Sequence: 1 MKVLAALVFLAGCQAKV.....AREGRLSIRRLGPLV 203

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Delop 6.0, Fgapext 7.0

Delop 6.0, Delet 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODL=frame+2n.model -DEV=xlp

-Q/cgn2.1/USPTO/US0982785/Runat_11032003_101610_27495/app.query.fasta.1.3576

-DB-EST -QMT=fastap -SUFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNIT5-bits -STRAT=1 -END=-1 -MATRIX=biosum62 -TRANS-human4.0.cdi -LIST=45

-DOCALIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US0982785.dcgcn.1.1.8826_gunat.11032003_101610_27495 -NCPU=6 -ICPU=3

-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120

-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estlin:*

4: em_estlmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_estl2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1014	100.0	690	13	B1599793
2	1014	100.0	699	12	B6769455
3	1014	100.0	748	12	B6706913
4	1014	100.0	748	13	B1553085
5	1014	100.0	757	13	B1603658
6	1014	100.0	757	13	BM042228
7	1014	100.0	790	13	B1551811
8	1014	100.0	798	12	B6708414
9	1014	100.0	800	13	BM042094
10	1014	100.0	817	12	B6774871
11	1014	100.0	845	12	B6829472
12	1014	100.0	933	13	B1668318
13	1014	100.0	965	14	B0677266
14	1011	99.7	692	13	B1546740
15	1010	99.6	802	13	B1458355
16	1010	99.6	811	13	B1600563
17	1010	99.6	942	12	B1600906
18	1009	99.5	797	12	B6715366
19	1009	99.5	808	13	B1668329
20	1007	99.3	741	12	B6762924
21	1006	99.2	725	12	B6708285
22	1006	99.2	748	12	B6715607
23	1006	99.2	803	13	B1670350
24	1006	99.2	842	12	B6763371
25	1006	99.2	891	13	B1548292
26	1006	99.2	1027	13	B1670367
27	1005	99.1	790	12	B6707147
28	1005	99.1	938	12	B6761746
29	1004	99.0	677	13	B1604042
30	1004	99.0	790	13	B1551066
31	1002	98.8	1100	14	BM914382
32	1001	98.7	687	12	B6706243
33	1001	98.7	693	12	B6707355
34	1001	98.7	812	13	B1601551
35	1000	98.6	856	13	BM042032
36	999	98.5	1030	14	B0682454
37	998	98.4	705	12	B6715556
38	998	98.4	919	13	B1551475
39	997	98.3	907	12	B6706129
40	993	97.9	941	14	B0681329
41	992	97.8	680	10	AV702686
42	989	97.5	1023	14	B0680425
43	988	97.4	669	10	AV655989
44	987	97.3	716	12	B6703146
45	987	97.3	757	12	B6703504

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
B1599793	B1599793	B1599793	B1599793	B1599793	B1599793	B1599793	690 bp	mRNA	EST 07-SEP-2001	
		603245228F1 NIH_MGC_96 Homo sapiens								
		CDNA clone IMAGE:5287710 5',								
		linear								
		EST								
		NIH-MGC http://mgc.nci.nih.gov/								
		National Institutes of Health, Mammalian Gene Collection (MGC)								

**JOURNAL
COMMENT**

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@rsf@u.washington.edu
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihai
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLAM11727 row: g column: 07
high quality sequence stop: 686.

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5287710"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtagcag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carlnici, in preparation). Library
constructed by M. Brownstein (NIH/NHRI, National
Institutes of Health). Note: this is a NIH_MGC library."

```

Alignment Scores:

Pred. NO.:	1,78e-105	Length:	69
Score:	1014.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-09-827-854-15_COPY_1_203 (1-203) x BI599793 (1-690)

QY	1	MettysValleTrrPALalaleuLeuValTrrPhelLeuAlaGlyCysGlnAlaLysVal	20
Db	73	ATGAAGGTTCTGTGGGCTGGCTGCTGGTCAATTCTTGGCAGGATGCCAGGCCAAGGTG	132
QY	21	GIuGlnAlaValaGIuThrGIuProGIuProGIuLeuAArgGlnGlnAlaThrGIuTrpGlnSer	40
Db	133	GAGCAAGCGGTGGAGACACAGAGCGGAGCCGACAGCTGGCCGACACAGACAGTGGCAGAGC	192
QY	41	GlyGlnAArgTrrPGLuLeuAlaLeuGlnLyaRphetrPasPrrLeuAArgTrrValGIuThr	60
Db	193	GGCCAGCCCTGGGAACCTGGCACTGGGTGCTTTGGGATTACCTCGCGCTGGGTGCAGACA	252
QY	61	LeuSerGIuGlnValGlnGlnGlnLeuLeuSerSerGlnAlaThrGlnGlnLeuAArgAla	80
Db	253	CTGTCTGGACGAGTGCACGAGAGCTGCTCAAGTCCAGGTACCACAGAACTGAGGGCG	312
QY	81	LeuMetAspGIuThrMetLysGlnLeuLeuLysAlaTrrLysSerGlnLeuGlnGlnLeu	100
Db	313	CTGATGGACGAGACCATTAAGAGATTGAAGGCTTCAAAATCGGAATGGAGGAACAACTG	372
QY	101	ThrProValAlaGlnGlnThrArgAlaAArgLeuSerLysGlnLeuGlnAlaAlaGlnAla	120
Db	373	ACCCGGGTGGGAGGAGAACGGGGCAGCGCTGTCCAAAGAGACTGCAGCGGGCGCAGGCC	432
QY	121	ArgLeuGlnLyaAspMetGlnAspValaCysGlnLyaArgLeuValGlnTrrArgGlnGlnVal	140
Db	433	CGGCTGGGCGCGGACATGAGGAGACTGTGTCGGGCCCTGTGTCACTACCGCGCGCAGGTG	492
QY	141	GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuAArgValAArgLeuAlaSerHisLeuAArg	160

Db	493	CAGGCCATGCTGCTGGGCACAGACACGAGAGCTGGGGGTGCGGCTCGGCTCCACCTGGGC	552
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	553	AAGCTGCGCTAAGCGGCTCCTCCGCGATGCGCGATGACCTGCACAAACGCGCTGCAGCTTAC	612
QY	181	GlnAlaGlyAlaArgGlnGlnValaLeuArgGlyLeuSerAlaIleArgLnuArgLeuGly	200
Db	613	CAGGCGGGGGCCCCGAGAGGCGCCGAGGCGGCGCTCACGCGCCATCCGCGAGCGCCTGGGG	672
QY	201	ProLeuVal	203
Db	673	CCCTCTGGTG	681

RESULT 2

LOCUS	699 bp	mRNA	linear	EST 15-MAY-2001
DEFINITION	602745580P1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4872427 , mRNA sequence.			

ACCESSION	BG/69455
VERSION	BG769455.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
1 (bases 1 to 699)
Eukaryota; Metazoa; Choriata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

FEATURES

```

source
1. .699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4872427"
/clone_1id="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCCGCGAG(c). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. !"

```

Alignment Scores:

Pred. No.:	1.82e-105
Score:	1014.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	12
	Gaps: 0
	Matches: 2030
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BG769455 (1-699)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

[illegible][illegible]

LOCUS B1553085 748 bp mRNA linear EST 05-SEP-2001
 DEFINITION 603196058p1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5275358 5',
 mRNA sequence.
 ACCESSION B1553085
 VERSION B1553085.1 GI:15440397
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 CHORDATA; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 748)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshitsuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11695 row: d column: 15
 High quality sequence stop: 748.
 Location/Qualifiers
 1..748
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5275358"
 /clone_lib="NIH_MGC_95"
 /tissue_type="hippocampus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI, Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.5 kb and
 normalized to R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 136 a 227 c 288 g 97 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2e-105 Length: 748
 Score: 1014.00 Matches: 203
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0
 US-09-827-854-15_COPY_1_203 (1-203) x B1553085 (1-748)
 QY 1 MettysValleuTRPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
 |||||||
 DB 75 ATGAAGGTTCTGTGGCTGCTGTCGTCACATTCCTGCGAGGATGCCAGGCCAAGG 134
 |||||||
 QY 21 GtGlnAlaValAlaGluThrGluProGluProGluLeuValArgGlnThrGluTrpGlnSer 40
 |||||||
 DB 135 GAGCAAGCGGTGAGACAGAGCCGAGCCGAGCTGCCACAGACCCAGTGGCAGAGC 194
 |||||||
 QY 41 GtGlnATGTTGCTGLeuAlaLeuGlyArgPheTrpAspTyrLeuValArgTrpValGlnThr 60
 |||||||
 DB 135 GCCCGACCGCTGGGAAGTGGCTGCTTTGGGATTACCTGGCTGGGCTGCAGACA 254
 |||||||
 QY 61 LeuSerGtGlnAlaValGtGlnGluLeuLeuSerSerGlnValThrGtGlnLeuValArgAla 80
 |||||||
 DB 225 CTGCTGTGAGACAGTGCAGAGAGAGCTGCACCTCCACAGTCAACCAGGAAGTGAAGGGG 314
 |||||||
 QY 81 LeuMetAspGluThrMetLeuGluLeuValArgTyrLysSerGtGluLeuGtGlnGluLeu 100
 |||||||

|||||
 DB 315 CTGATGAGACAGACCATGAGAGAGGCTCAAAATCGGAACGTGAGGAACAACATG 374
 |||||||
 QY 101 ThrProValAlaGtGluThrArgAlaArgLeuSerTyrGtGluLeuGlnAlaAla 120
 |||||||
 DB 375 ACCCGGTGCGGAGAGACGCGGCGACGCTGTCTCAAGGACCTGCAGCGCGCAGGGC 434
 |||||||
 QY 121 ArgLeuGtAlaAspMetGtGlnAspValCysGtArgLeuValGtThrArgGtGluVal 140
 |||||||
 DB 435 CGGCTGGCGCGACATGAGAGACGTCGCGCCCTGTCAGTACCGCGCGCAGGTCG 494
 |||||||
 QY 141 GtAlaMetLeuGtGlnSerThrGtGluLeuValArgLeuAlaSerTyrValArg 160
 |||||||
 DB 495 CAGGCCATCTCTGGCCAGACACCGAGAGCTGGGCTGCTGCCCTCCACCTGGCC 554
 |||||||
 QY 161 LysLeuArgLysArgLeuValArgAspAlaAspLeuGtGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 555 AAGCTGCGTAACGCGCTCTCCGCCATGCCATGACCTCAACAGCGCTGCAGTATC 614
 |||||||
 QY 181 GtAlaGtAlaArgGtGluGtAlaGtGluArgGtLysSerAlaTleArgGtGluArgGtGly 200
 |||||||
 DB 615 CAGCGCGGCGCGCGAGGCGCGCGAGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 674
 |||||||
 QY 201 ProLeuVal 203
 |||||||
 DB 675 CCCCTGGTG 683
 |||||||
 RESULT 5
 B1603658
 LOCUS B1603658 757 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603251982p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5303396 5',
 mRNA sequence.
 ACCESSION B1603658
 VERSION B1603658.1 GI:15496597
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 CHORDATA; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 757)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshitsuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11768 row: d column: 21
 High quality sequence stop: 753.
 Location/Qualifiers
 1..757
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5303396"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site 1: BamHI, Site 2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.3 kb and
 normalized to R0T 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 138 a 230 c 292 g 97 t

ORIGIN

Alignment Scores:

Pred. No.: 2,03e-105
 Score: 1014.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%

DB: 13
 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x B1603658 (1-757)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 73 ATGAAGGTTCTGTGGCGCTGCTGGTGCATTCCTGGCAGATGCGAGATGCCAAGGTTG 132
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 133 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTCCGACAGACAGACGAGTGGCAGAGC 192
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 193 GGCCAGCGCTGGAGAACCTGGCACTGGCTGCTTTTGGAATTAACCTGCGCTGGTGCAGACA 252
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 253 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAAGTCCCAAGTCCAGAGAACTGAGGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
 DB 313 CTGATGAGCAGAGACCATGAGAGAGTTGAAGGCTTCAAAATCGGAAGTGGAGAACACTG 372
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 373 ACCCGGTGGGGAGAGACAGCGGGCAGCGTGTCCAAGAGACTGCAGCGCGCAGGGCC 432
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 DB 433 CGGCTGGGGCGGAGATGAGAGACCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 492
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 493 CAGGCGCATGCTCGCGCAGACAGCACCGAGAGCTGCGGCTGCGCTCCACACTGCGGC 552
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 553 AAGCTGCGTAAAGCGGCTCTCCGCGATGCCATGACCTGCAGAAAGCGCTGCACTGTAC 612
 QY 181 GlnAlaGlyAlaArgGluGlyAlaGlnArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 613 CAGCGCGGGCGCGGAGAGGGCGCCAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 672
 QY 201 ProLeuVal 203
 DB 673 CCCCTGTGTG 681
 RESULT 6
 LOCUS BM042228 757 bp mRNA linear EST 07-NOV-2001
 DEFINITION 603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
 mRNA sequence.
 ACCESSION BM042228
 VERSION BM042228.1 GI:16771495
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 757)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.

FEATURES

source

Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: LNCM2021 row: k column: 21
 High quality sequence stop: 757.

Location/Qualifiers

1..757

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5557004"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 SuperScript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 134 a 229 c 296 g 98 t
 ORIGIN

Alignment Scores:

Pred. No.: 2.03e-105
 Score: 1014.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%

DB: 13
 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BM042228 (1-757)

QY 1 MetLysValLeuTrpAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 50 ATGAAGGTTCTGTGGCGCTGCTGGTGCATTCCTGGCAGATGCGAGATGCCAAGGTTG 109
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 110 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTCCGACAGACAGACGAGTGGCAGAGC 169
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 170 GGCCAGCGCTGGAGAACCTGGCACTGGCTGCTTTTGGAATTAACCTGCGCTGGTGCAGACA 229
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 230 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAAGTCCCAAGTCCAGAGAACTGAGGGCG 289
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu 100
 DB 290 CTGATGAGCAGAGACCATTAAGAGAGTTGAAGGCTTCAAAATCGGAAGTGGAGAACACTG 349
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 350 ACCCGGTGGGGAGAGACAGCGGGCAGCGTGTCCAAGAGACTGCAGCGCGCGCAGGGCC 409
 QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 DB 410 CGGCTGGGGCGGAGATGAGAGAGCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 469
 QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 470 CAGGCGCATGCTCGCGCAGACAGACGAGAGCTGCGGGTGGCTGCTCCACACTGGCGC 529
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180

DB 530 AACCTGCTAAGCGGCTCCTCCGCGATGCCGATGACCGAGAAGCGCTGCGATGTAC 589

QY 181 GlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaIleArgGluArgGly 200
|||||
DB 590 CAGCGCGGCGCCGCGAGCGCGGCTGACGCGCATCCGCGAGCGCTGGG 649

QY 201 ProLeuVal 203
|||||
DB 650 CCCCTGGTG 658

RESULT 7
BI551811 790 bp mRNA linear EST 05-SEP-2001
LOCUS 603192716F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:526415 5',
DEFINITION mRNA sequence.
ACCESSION BI551811
VERSION BI551811.1 GI:15439123
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 790)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM1665 row: P column: 04
High quality sequence stop: 777.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:526415"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: Brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 141 a 235 c 311 g 103 t

ORIGIN

Alignment Scores:
Pred. No.: 2.16e-105 Length: 790
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x BI551811 (1-790)

QY 1 MellyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
DB 75 ATGAAGGTTCTGTGGGCTGGTGGTCACATCTCTGCGAGATGCCAGGCCAAGGTG 134

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40

DB 135 GAGCAAGCGGTGAGACAGAGCGGAGCGGACCTCGCCAGACAGCCAGTGGCAGAC 194
|||||

QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspPylLeuArgTrpValGlnTr 60
|||||

DB 195 GGCACCGCTGGGAACTGGCTGGGTGGTTTGGATTACTGGCTGGCTGGGAGACA 254

QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||

DB 255 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGGTCCAGACCCAGAACTGAGGGGG 314

QY 81 LeuMetAspGluThrMetLeuGlyGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
|||||

DB 315 CTGATGAGACAGACCATGAAGGAGTGAAGGCCCTACAAATCCGAACTGGAGGAACTAG 374

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerGluLeuGlnAlaGlnAla 120
|||||

DB 375 ACCCGGTGGCGGAGAGAGCGGCGACGCTGTCCAGAGAGCTGCAGGCGCGCAGGCC 434

QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
|||||

DB 435 CGGCTGGCGCGGACATGAGAGAGTGTGGCGCGCTGTGCAGTACCGCGCGAGAGTG 494

QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||

DB 495 CAGGCCATGCTCGCGCAGAGACCGAGAGACTCGGGGTGGCTCCCTCCACCTGCGCG 554

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
|||||

DB 555 AACCTGCTAAGCGGCTCCTCCCGCATGCCGATGACTGCAGAAAGCGCTGGCAGGTAC 614

QY 181 GlnAlaGlyAlaArgGluGlyAlaArgGlyLeuSerAlaIleArgGluArgGly 200
|||||

DB 615 CAGCGCGGCGCCGCGAGCGCGCGGAGCGGCTCAGCGCCATCCGCGAGCGCTGGG 674

QY 201 ProLeuVal 203
|||||

DB 675 CCCCTGGTG 683

RESULT 8
BG708414 798 bp mRNA linear EST 07-MAY-2001
LOCUS 602672374F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794955 5',
DEFINITION mRNA sequence.
ACCESSION BG708414
VERSION BG708414.1 GI:13985733
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 798)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM10677 row: C column: 20
High quality sequence stop: 778.
Location/Qualifiers
1..798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4794955"
/clone_lib="NIH_MGC_96"


```

/issue_type="hypothalamus"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTGTGTTTGTGTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROP 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      143 a      235 c      319 g      101 t
ORIGIN

Alignment Scores:
Pred. No.:      2,19e-105      Length:      798
Score:          1014.00      Matches:      203
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              12      Gaps:      0

US-09-827-854-15_COPY_1_203 (1-203) x BG708414 (1-798)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      73 ATGAAGGTTCTGTGGGCTCGTTGCTGTCACATTCCTGGCAGATGCCAGCCAAAGTG 132
QY      21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      |||||||
Db      133 GAGCAAGCGGTGGAGACAGAGCGGAGCCGAGCTCGCAGACAGACCGAGTGCGCAGAGC 192
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
      |||||||
Db      193 GGCACACGCTGGAGACTGGCTGGCTTTGGATTAAGTGGGAGGAGGAGGAGGAGGAGGAG 252
QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||||
Db      253 CTGTCTGACAGAGTCAGAGAGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
      |||||||
Db      313 CTGATGACAGACACCTGGAAGAGGTTGAAGGCTTACAAATCGAATCGAGAGAACACATG 372
QY      101 ThrProValAlaGluLeuThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      |||||||
Db      373 ACCCGCGTGGCGAGAGACCGCGGACGCTCTCCAAAGAGCTGCAAGCGCGCGAGGCC 432
QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
      |||||||
Db      433 CGGCTGGCGCGGACATGAGAGGACGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 492
QY      141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      |||||||
Db      493 CAGGCCATGCTGCGGCGAGAGACCGGAGGAGTGGGCTGCGCTCCCAACCTGCCGCC 552
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      |||||||
Db      553 AACCTCGTAAAGCGGCTCTCCGCGATGCCGATGACCTCAGAGAGCGCTGCCAGGTGTA 612
QY      181 GlnAlaGlyAlaArgGluGluAlaGluLysArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      |||||||
Db      613 CAGGCGGGGGCGCGGAGGCGCGGCGGCTCAGCGCATCCGCGAGGCGCGCTCGGGG 672
QY      201 ProLeuVal 203
      |||||||
Db      673 CCCCTGGTG 681

RESULT 9      BM042094      800 bp      mRNA      linear      EST 07-NOV-2001
LOCUS      BM042094
DEFINITION  603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
      mRNA sequence.

```

```

ACCESSION      BM042094
VERSION        BM042094.1      GI:16771361
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/
AUTHORS        1 (bases 1 to 800)
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgaabp-r@mail.nih.gov
               Tissue Procurement: DCTD/DPF
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/ULNL at:
               http://image.llnl.gov
               Plate: L1CM1875 row: 1 column: 03
               High quality sequence stop: 792.
               Location/Qualifiers
               1..800
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5420618"
               /clone_lib="NIH_MGC_112"
               /issue_type="melanotic melanoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
               EcoRI; cDNA made by oligo-dt priming. Directionally cloned
               into EcoRI/XhoI sites using the following 5' adaptor:
               GGCAAGAG(G). Library constructed by Ling Hong in the
               laboratory of Gerald M. Rubin (University of California,
               Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
               Superscript II RT (Life Technologies). Note: this is a
               NIH_MGC Library."
BASE COUNT      146 a      239 c      315 g      100 t
ORIGIN

Alignment Scores:
Pred. No.:      2.2e-105      Length:      800
Score:          1014.00      Matches:      203
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              13      Gaps:      0

US-09-827-854-15_COPY_1_203 (1-203) x BM042094 (1-800)
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      50 ATGAAGGTTCTGTGGGCTCGTTGCTGTCACATTCCTGGCAGATGCCAGCCAAAGTG 109
QY      21 GluGlnAlaValAlaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      |||||||
Db      110 GAGCAAGCGGTGGAGACAGAGCCGAGCGGAGCTGCGCAGACAGACCGAGTGCGCAGAGC 169
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
      |||||||
Db      170 GGCACAGCGCTGGAGACTGGCTGCTTTGGATTAAGTGGGAGGAGGAGGAGGAGGAGGAG 229
QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||||
Db      230 CTGTCTGACAGAGTCAGAGAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
      |||||||
Db      290 CTGATGACAGACCATGAGAGAGTTGAAGGCTTACAAATCGAATCGAGAGAACACTG 349
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      |||||||

```

Db	Accession	Source	Organism	Reference	Title	Journal	Comment
Db	350	ACCCGGTGGCGGAGGAGACGGCGGCGACGGCTGTCCAAAGAGCTGACGGCGCAGGCC	409				
Qy	121	ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140				
Db	410	CGCGTGGGCGCGGACATGAGGACGTGTGCGGCCCTGTGTCAGTACCGCGCAGAGTG	469				
Qy	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerIleuArg	160				
Db	470	CAGGCCATGCTGGCGCAGACACCGAGGAGGTGCGGGGTGCGCTGCTCCACCTGGCGC	529				
Qy	161	IysIeArgIysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180				
Db	530	AAGCTGCGGAAACGGCTCTCCGCGATGCGGATGACCTCAGAAAGCGCTGGCAGTTC	589				
Qy	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly	200				
Db	590	CAGGCGGGGCGCGGAGGGCGCCGAGCGCGGCTTCAGCGCCATCGCGGAGGCGCTGGGG	649				
Qy	201	ProLeuVal	203				
Db	650	CCCCGTGTG	658				
RESULT	10						
LOCUS	BG774871						
DEFINITION	60264997551 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:4760827 5',						
ACCESSION	BG774871						
VERSION	BG774871.1						
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
TITLE	1 (bases 1 to 817)						
JOURNAL	NIH-MGC http://mhc.nci.nih.gov/ .						
COMMENT	Unpublished (1999)						
	Contact: Robert Strausberg, Ph.D.						
	Email: rsrausb@mail.nih.gov						
	Tissue Procurement: DCM/D/DP						
	CDNA Library Preparation: Ling Hong/Rubin Laboratory						
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)						
	DNA Sequencing by: Incyte Genomics, Inc.						
	Clone distribution: MGC clone distribution information can be						
	found through the I.M.A.G.E. Consortium/LNL at:						
	http://image.llnl.gov						
	Plate: L10C1612 row: e column: 20						
	High quality sequence stop:813.						

FEATURES	source	Location/Qualifiers
		1. .817
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:4760827"
		/clone_lib="NIH_MGC_40"
		/tissue_type="carcinoma, cell line"
		/lab_host="DH10B (phage-resistant)"
		/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
		Site_2: EcoRI; cDNA made by oligo-dT priming.
		Directionally cloned into EcoRI/XhoI sites using the
		following 5' adaptor: GGCACAG(G). Library constructed by
		Ling Hong in the laboratory of Gerald M. Rubin (University
		of California, Berkeley) using Zap-cDNA synthesis kit
		(Stratagene) and Superscript II RT (Life Technologies).
		Note: This is a NIH_MGC Library."
BASE COUNT	148 a	242 c 324 g 103 t
ORIGIN		

Alignment Scores:	2.26e-105	817
Pred. No.:	1014.00	Matches:
Score:	100.00%	Conservative:
Percent Similarity:	100.00%	Mismatches:
Best Local Similarity:	100.00%	0

Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0
US-09-827-854-15_COPY_1_203 (1-203) x BG774871 (1-817)			
QY	1 MettysValleuTrrpAlaAlaLeuLeuValTthrPheLeuAlaGlyCysGlnAlaIysVal	20	
DB	40 ATGAAGATTCTGTGGGCTCGTGGTGGTACATTCTCGAGAGATCCAGGCGCAAGTG	99	
QY	21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThruTrpGlnSer	40	
DB	100 GAGCAAGCGGTGGAGACAAAGCCGGAGCCCGAGCTCGCCAGACAGCCAGTGGCGAGAC	159	
QY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60	
DB	160 GGCCAGCGCTGGGAATGTCACCTGGGTCCCTTTGGATTACCTGGCTGGGGTGGAGAA	219	
QY	61 LeuSerGluGlnValaGlnGlnGluLeuLeuSerSerGlnValaThrGlnGluLeuArgAla	80	
DB	220 CAGTCTGACAGAGTGCAGGAGGAGGAGCTGCTCACCTCCACAGTCAACCCAGAACTGAGGCG	279	
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu	100	
DB	280 CGATGTGAGAGAACCATGATGAAGATTGAAGGGCTCAAAATCGGAATGGAGAAACAACTG	339	
QY	101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnGlnAlaAlaGlnAla	120	
DB	340 ACCCGCGTGGCGGAGAGACGGCGGCGACGGCTGTCCAAAGAGCTGAGGCGGCGCAGGCG	399	
QY	121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal	140	
DB	400 CGGCTGGGCGGCACATGAGGAGCGTGTGGCGCCGCTGTGCAGTACCGCGCGAGGTG	459	
QY	141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160	
DB	460 CAGGCGATGCTGTGGCCAGAGCACCGAGGAGCTGGCGGTGGCTGCCCTCCACCTGGCCC	519	
QY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp	180	
DB	520 AAGCTGCCGTAACCGGCTCCTCCGCATGCGCATGACCTCCAGAAAGCGCTGGCAGGTAC	579	
QY	181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGly	200	
DB	580 CAGGCGGCGGCGCGGAGGCGCCGAGCGCTCAGCGCCATCCGCGAGGCGCTGGG	639	
QY	201 ProLeuVal	203	
DB	640 CCCCTGGTG	648	

RESULT 11
BG829472
LOCUS
DEFINITION BG829472 845 bp mRNA linear EST 22-MAY-2001
60726376881 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:4892112 5',
mRNA sequence:
ACCESSION BG829472
VERSION BG829472
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 845)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1995)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

CNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (tLNU)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MCC clone distribution information can be
 found through the I.M.A.G.E. Consortium/tLNU at:

QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 255 CTGTCTGAGCAGGTGCGAGGAGCTGCTCAGCTCCAGGTACACCGAGAACTGAGGGCG 314
QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
|||||
Db 315 CTGATGACAGACCAATGAGAGATTGAAAGCCTACAAATCGGAACCTGGAGAACTG 374
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGlnAlaAlaGlnAla 120
|||||
Db 375 ACCCGCGTGGCGAGAGAGACCGCGCAGCTGTCCAGAGAGCTGCAGGGGGCGAGGCC 434
QY 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGlnTyrArgGIuVal 140
|||||
Db 435 CGGCTGGCGCCGACATGAGAGACGTGTGGCGCGCTGTGACAGTACCGCGCGAGGTG 494
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValAlaArgLeuAlaSerHisLeuArg 160
|||||
Db 495 CAGGCGATGCTCGCGACAGACCGAGAGCTCGGGGTGGCTCGCTCCACCTGGCGC 554
QY 161 LysLeuArgTyrArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
|||||
Db 555 AAGCTGCTAAGCGGCTCCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 614
QY 181 GlnAlaGIuAlaArgGIuGIuAlaArgGIuLysSerAlaIleArgGIuArgLeuGIu 200
|||||
Db 615 CAGCGCGGGCGCGGAGCGCGGAGCGCGCTCAGCGCCATCCGCGAGCGCCGTGGGG 674
QY 201 ProLeuVal 203
|||||
Db 675 CCCCTGGTG 683

RESULT 13
LOCUS B0677266 965 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8494742 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6295132
5', mRNA sequence.
ACCESSION B0677266
VERSION B0677266.1 GI:21789945
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 965)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2500 row: 0 column: 05
High quality sequence stop: 593.
Location/Qualifiers
1..965

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6295132"
/clone_lib="NIH_MGC_112"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH-MGC Library."

BASE COUNT 172 a 328 c 326 g 139 t

ALIGNMENT SCORES:

Pred. No.: 2,86e-105 Length: 965
Score: 1014.00 Matches: 203
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-827-854-15_COPY_1_203 (1-203) x B0677266 (1-965)

QY 1 MetLysValLeuThrPAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||
Db 54 ATGAAAGTTCTGTGGCTGCTGCTGTGTACATTCCTGGCAGGATGCCAGGCCAAGTG 113
QY 21 GluGlnAlaValGIuThrGIuProGIuProGIuLeuArgGlnGlnThrGIuTrpGlnSer 40
|||||
Db 114 GAGCAACGCGGTGAGACAGACGCGGAGCCGAGCTCGCCAGAGACGAGTGGCAGAGC 173
QY 41 GlyGlnArgTrpGIuLeuAlaLeuGIuLysArgPheTrpAspTyrLeuArgTrpValGlnThr 60
|||||
Db 174 GGCACAGCGCTGGGAACCTGACAGCTGCTGCTTTGGATTACCTCGCGTGGGTGCAGACA 233
QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
|||||
Db 234 CTGTCTGAGCAGGTGACAGAGAGAGCTGTACAGCTCCAGGTACCCAGGAACCTGAGGGCG 293
QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
|||||
Db 294 CTGATGAGAGAGACCAATGAGAGAGTTGAAAGCCCTCAAAATCGGAATGGAGAGAACAG 353
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGlnAlaAlaGlnAla 120
|||||
Db 354 ACCCGCGTGGCGGAGAGAGACCGCGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGAGGCC 413
QY 121 ArgLeuGIuAlaAspMetGIuAspValCysGIuArgLeuValGlnTyrArgGIuVal 140
|||||
Db 414 CGGCTGGCGCGGACATGAGAGAGAGCTGTGCGCGCTGTGACAGTACCGCGCGAGGGTG 473
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValAlaArgLeuAlaSerHisLeuArg 160
|||||
Db 474 CAGGCGATGCTCGCGCAGAGACCGGAGAGCTGGGGTGGCTGCCCTGCCACCTGGCGC 533
QY 161 LysLeuArgTyrArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
|||||
Db 534 AAGCTGCTAAGCGGCTCCTCCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 593
QY 181 GlnAlaGIuAlaArgGIuGIuAlaArgGIuLysSerAlaIleArgGIuArgLeuGIu 200
|||||
Db 594 CAGGCGGGGGCGCGAGGAGCGCGCGGCTGTACAGCCATCCGCGAGCGCTGGGG 653
QY 201 ProLeuVal 203
|||||
Db 654 CCCCTGGTG 662

RESULT 14
LOCUS B1546740 692 bp mRNA linear EST 05-SEP-2001
DEFINITION 60319104F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5262744 5',
mRNA sequence.
ACCESSION B1546740
VERSION B1546740.1 GI:15434052
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 692)


```
QY      1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
      |||||||
Db      75 ATGAAGTTCTGTGGGCTGTGGTGTGTCACATTCCTGGCAGGATGCCAGCCAAAGCTG 134
QY      21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
      |||||||
Db      135 GAGCAAGCGGTGGAGACAGACGCGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 194
QY      41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
      |||||||
Db      195 GGCACAGCGCTGGGAACCTGGACTGGGTCCGCTTTGGGATTACTGCGCTGGGTGCAGACA 254
QY      61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
      |||||||
Db      255 CTGTCTGACGACAGGTGACGAGAGCTGCTCAGCTCCAGGTCCACCGAGAACTGAGGGCG 314
QY      81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluGlnLeu 100
      |||||||
Db      315 CTGATGACGACGACCATGAAGAGTTGAAGGCTTACAAATCGGAACTGGAGAACAACTG 374
QY      101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
      |||||||
Db      375 ACCCGGTTGGCGGAGAGACGCGGCGACGGCTGTCCAAGAGCTGCAGGGCGCGCAGGCC 434
QY      121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
      |||||||
Db      435 CGGCTGGGCGCGACATGGAGACGTGTGCGGCGCGCTGGTGCAGTACC GCGCGGAGGTG 494
QY      141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
      |||||||
Db      495 CAGGCCATGCTCGGCCAGAGACACCGAGAGCTGCGGGTGGCTCGCTCCCACTGGCGC 554
QY      161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
      |||||||
Db      555 AAGCTGCGTAAGGGGCTCTCCCGCATGCCGATGACCTGCAGAGCGCCTGGCAGTGTAC 614
QY      181 GlnAlaGlyAlaArgGluGluAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
      |||||||
Db      615 CAGGCCGCGGGCGCGAGGGCGCGAGCGCGGCTCAGGCCATCCGCGAGCGCGCTGGGG 674
QY      201 ProLeuVal 203
      |||||||
Db      675 CCCCTGGTG 683
```

Search completed: March 14, 2003, 20:13:59
Job time : 829.891 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:08:57 ; Search time 2016.1 Seconds
(without alignments)
3998.551 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Perfect score: 1374
Sequence: 1 MKVLMAALVFLAGCOAKV.....RAKLEBOAQOIRLQAEAFQA 277

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=rlp
-Q=/cgn2_1/USPTO.spool/US09827854.rgnat_11032003_101610_27486/app.query.fasta_1.3576
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=psio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09827854.6cgn_1.1.13008_ggnat_11032003_101610_27486 -NCPU=6 -ICPU=3
-NO_XLPEXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOP -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: GenEmbl:*
2: gb_ba:*
3: gb_hlg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_sts:*
13: gb_sy:*
14: gb_un:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_in:*
20: em_mu:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pal:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	1110	6 E00359	E00359 CDNA coding
2	1374	100.0	1110	6 E00823	E00823 DNA sequence
3	1374	100.0	1147	6 AX302545	AX302545 Sequence
4	1374	100.0	1156	6 BD004278	BD004278 Apo E hum
5	1374	100.0	1156	6 BD004278	BD004278 Apo E hum
6	1374	100.0	1186	6 BC003557	BC003557 Homo sapi
7	1367	99.5	1110	6 E08423	E08423 DNA coding
8	1365	99.4	1156	6 BD004277	BD004277 Apo E hum
9	1365	99.3	1157	6 AX333278	AX333278 Sequence
10	1365	99.3	1157	6 AX409597	AX409597 Sequence
11	1365	99.3	1157	6 H15975	H15975 Sequence 1
12	1365	99.3	1157	6 H15975	H15975 Sequence 1
13	1362	99.1	1156	6 BD004279	BD004279 Apo E hum
14	1352	99.1	1178	6 MFAPOE	M12529 Human apoli
15	1185.5	86.3	5491	6 AF261279	AF261279 Monkey mRNA
16	1185.5	86.3	5491	6 AF261279	AF261279 Homo sapi
17	1185.5	86.3	41907	6 AF050154	AF050154 Homo sapi
18	1185.5	86.3	107567	9 AC011481	AC011481 Homo sapi
19	1185.5	86.3	5515	9 HUMAPOE4	M10065 Human apoli
20	1142	83.1	5413	9 AF261280	AF261280 Pan trogl
21	1079	78.5	1138	4 AF303830	AF303830 Typaia gl
22	1067	77.7	208239	2 AC021988	AC021988 Homo sapi
23	1065	77.5	4762	9 BABAPOE	M29322 Baboon apoli
24	1021.5	74.3	1060	4 BABAPOE	M36603 Rabbit apoli
25	1015	73.9	1045	10 MUSAPOE	M12414 Mouse apoli
26	1015	73.9	1104	10 BC028816	BC028816 Mus muscu
27	1004.5	73.1	965	6 AX384545	AX384545 Sequence
28	1004.5	73.1	1108	4 BPAPOE	X61171 B. taurus mr
29	1004.5	73.1	5617	6 AX384541	AX384541 Sequence
30	1004.5	73.1	6026	6 AX384539	AX384539 Sequence
31	1002.5	73.0	1126	6 AR164342	AR164342 Sequence
32	1002.5	73.0	1126	6 AR205885	AR205885 Sequence
33	996.5	72.5	1122	4 SSAPOE	X72835 S. scrofa mr
34	986.5	71.8	1154	4 BPAPOE	X64839 B. taurus mr
35	984	71.6	660	6 A62340	A62340 Sequence 28
36	984	71.6	660	6 A62342	A62342 Sequence 30
37	984	71.6	660	6 A62344	A62344 Sequence 32
38	984	71.6	660	6 AR075563	AR075563 Sequence
39	960	69.9	951	10 S76779	S76779 rAPOE-apoli
40	943	68.6	718	9 AF20049753	AF200499 Pan trogl
41	939	68.3	718	9 AF20050053	AF200502 Gorilla g
42	938	68.3	959	10 MUSAPOE	M73490 Mus musculu
43	936	68.1	718	9 AF20050353	AF200505 Pongo pyg
44	932	67.8	718	9 AF20050653	AF200508 Hylobates
45	922.5	67.1	1069	10 RATAPOE	J00705 rat apolipo

RESULT 1

ALIGNMENTS

E00359 LOCUS E00359 1110 bp RNA linear PAT 29-SEP-1997
 DEFINITION cDNA coding human apolipoprotein E3.
 ACCESSION E00359
 VERSION E00359.1 GI:2168646
 KEYWORDS JP 1985118189-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1110)
 Teranishi, Y., Takamatsu, N., Matsui, Y., Kimura, M. and Ikeda, Y.
 REFERENCE DNA FRAGMENT
 TITLE Patent: JP 1985118189-A 1 25-JUN-1985;
 JOURNAL MITSUBISHI CHEM IND LTD
 COMMENT OS human
 PN JP 1985118189-A/1
 PD 25-JUN-1985
 PF 29-NOV-1983 JP 1983224980
 PI TERANISHI YUTAKA, TAKAMATSU NOBUHIKO, MATSUI YASUSHI, PI
 KIMURA MASAKO,
 PI IKEDA YASUKO
 PC C12N15/00, C07H21/04, C12P21/00;
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue; location=livr;
 FH Key Location/Qualifiers
 FT CDS 15..968
 FT sig_peptide 15..68
 FT mat_peptide 69..965
 FT /product='apolipoprotein E3 signal peptide' FT
 FT /product='apolipoprotein E3'.
 FEATURES
 source Location/Qualifiers
 1..1110
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 198 a 353 c 416 g 143 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-72 Length: 1110
 Score: 1374.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-827-854-15_COPY_1_277 (1-277) x E00359 (1-1110)
 QY 1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db |||||||
 Db 15 ATGAAGTTCTGTGGCTGCTTCCTGTCACATTCCTGCAGATGCCAGCCAAAGTG 74
 QY 21 GluGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 Db |||||||
 Db 75 GAGCAACGCGTGGAGACAGACCCGAGCTGCGCCAGCAGACCGAGTGCACAGC 134
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 Db |||||||
 Db 135 GGCAGCGCTGGAGAACTGCGACTGCTGCTTTGGGATTAACCTGCCCTGGGTGCAGACA 194
 QY 61 LeuSerLeuGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db |||||||
 Db 195 CTGCTTAGCAGAGTGCAGAGGAGACTCTCAGCTCCAGGTCCAGCAGGAAGTGAAGGCG 254
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 Db |||||||
 Db 255 CTGATGACGACGACCATGAAGGAGTTGAAGGCTTACAAATCGGAATCGAGAACTG 314

QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 Db |||||||
 Db 315 ACCCGCTGGCGGAGGAGAGCGCGGCGACGCTGCCAAGAGCTGCAGCGCGGACGCC 374
 QY 121 ArgLeuGlnValAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyVal 140
 Db |||||||
 Db 375 CGGCTGGCGCGGACATGAGAGACGTGTGGCGCGCTGTGTGACGTACCCCGGAGAGTG 434
 QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db |||||||
 Db 435 CAGGCATGCTCGGCGCAGAGACAGCAGAGAGTGGGGTGGCGCTCCGCTCCACCTGCGC 494
 QY 161 LysLeuArgLysArgLeuLeuArgAlaAspAlaAspAspLeuLysArgLeuAlaValTyr 180
 Db |||||||
 Db 495 AAGCTGGTAAAGCGGCTCCCTCCGATGCCATGACCTGAGAAAGCCCTGGCAGTGTAC 554
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaArgGlyLeuSerLysIleArgGluArgLeuGly 200
 Db |||||||
 Db 555 CAGGCGGGGCGCGCGAGGCGCGGAGCGGCGCTCAGGCCATTCGCGAGCGCTGGGG 614
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 Db |||||||
 Db 615 CCCTGTGTGAACAGGCGCGCTGCGGCGCGCCACTGTGTGGTCTCCCTGCGCGGCGAGCGC 674
 QY 221 LeuGlnGluArgAlaGlnAlaTTPGlyGluArgLeuArgAlaArgMetGluMetGly 240
 Db |||||||
 Db 675 CTACAGGAGCGGCGCGAGCTGGGGGAGCGGCTGCGCGCGGCGGATGAGAGATGGGC 734
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
 Db |||||||
 Db 735 AGCCGAGCCCGGACCGCTCGAGCAGAGTGAAGAGAGAGTGGCGGAGTGGCGGCAAG 794
 QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleGlnAla 277
 Db |||||||
 Db 795 CTGAGAGAGCAGGCGCCAGCATATCGCTGAGCGCGAGCTTCAGGCC 845
 RESULT 2
 E00823 1110 bp DNA linear PAT 29-SEP-1997
 LOCUS E00823
 DEFINITION DNA sequence coding for human apolipoprotein E and its signal peptide.
 ACCESSION E00823
 VERSION E00823.1 GI:2169084
 KEYWORDS JP 1986096997-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1110)
 AUTHORS Teranishi, Y., Matsui, Y., Ikeda, Y. and Kimura, M.
 TITLE PRODUCTION OF HUMAN APOLIPOPROTEIN E-LIKE PROTEIN
 JOURNAL Patent: JP 1986096997-A 1 15-MAY-1986;
 MITSUBISHI CHEM IND LTD
 COMMENT OS Human (Homo sapiens)
 PN JP 1986096997-A/1
 PD 15-MAY-1986
 PF 16-OCT-1984 JP 1984216987
 PI TERANISHI YUTAKA, MATSUI YASUSHI, IKEDA YASUKO, KIMURA MASAKO
 PC C12P21/00, A61K35/74, A61K37/04, C12N15/00, (C12P21/00, C12R1:19),
 PC (C12N15/00,
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: tissue; location=livr;
 FH Key Location/Qualifiers
 FT 3'UTR 1..14
 FT sig_peptide 15..68
 FT /product='human apolipoprotein E signal peptide' FT
 FT CDS 69..968
 FT /product='human apolipoprotein E' FT 3'UTR

[illegible][illegible]

QY	201	ProlevalGlucInglnglyrvalArpAlaIatThrAlaGlySerLeuAlaGlyInPro	220
Db	646	CCCTGCTGTGAACAGAGGCGCGCTGGCGGCGCCACTGTGGCTCCCTGGCGGCGGCGG	705
QY	221	LeuGIngluArAlaGInaIatRptGlyGluArGleuArGlaArGMeGluGluMetGly	240
Db	706	CTACAGACAGCGGGCCAGGCGCTGGGGCGAGCGGCTGGCGCGGATGAGAGATGGGGC	765
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	766	AACCCGACCCCGGACCGCTCGACAGAGGAGGAGAGCGAGTGGCCGAGTGGCGCCCAAG	825
QY	261	LeuGIngluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla	277
Db	826	CTGGAGGACGAGGCCACGACGATACGCTGCGAGCGAGGCGCTTCAGAGCC	876
RESULT 4			
LOCUS	BD004278	1156 bp	DNA
DEFINITION	Apo E humanized mammal.		Linear
ACCESSION	BD004278.1		
VERSION	BD004278.1 GI:18632239		
KEYWORDS	JP 2001017028-A/2.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	I (bases 1 to 1156)		
TITLE	Fujita,S., Hamanaka,H., Fukui,Y. and Yokoyama,M.		
JOURNAL	Apo E humanized mammal Patent: JP 2001017028-A 2 23-JAN-2001; MITSUBISHI CHEMICAL CORP		
COMMENT	OS Homo sapiens (human) PN JP 2001017028-A/2 PD 23-JAN-2001 PF 28-APR-2000 JP 2000128919 PR SHINOBU FUJITA,HIROKI HAMANAKA,YUKO FUKUI,MINESUKE YOKOYAMA PC A01K67/027,A51K45/00,A61P25/28,A61P43/00,C12N5/10,PC C12N15/09//C07K14/775, PC (C12N5/10,C12R1:91),C12N5/00,C12N15/00,(C12N5/00,C12R1:91) CC		
FEATURES			
source	FT Key Location/Qualifiers CDS Location/Qualifiers 1..1156 /organism="Homo sapiens" /db_xref="taxon:9606"		
BASE COUNT	208 a 368 c 432 g 148 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.26e-72	Length:	1156
Score:	1374.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-827-854-15_COPY_1-277 (1-277) x BD004278 (1-1156)			
QY	1	MellysValIleuTRPALaAlaLeuAlaValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	61	ATGAGGTTCTGTGGCGCGCTGTGTCATCTCCGCGCAGATGCGCAGGCCAAGTGG	120
QY	21	GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer	40
Db	121	GAGCAAGCGTGGAGACAGAGCGCGGACCGCAGCTGCCACAGACCGAGTGGCAGAGC	180
QY	41	GlyGlnArgTrpGluLeuAlaLeuGluLysArgPheThrPaspArgTrpLeuArgTrpValGlnThr	60
Db	181	GCCAGCCCTGGGAACCTGCACTGGTGGCTTTTGGATTACTTCGCTGGGTGCAGACA	240

OY	61	LeuSerGIUGInValGlnGluGluLeuLeuSerSeGIInValThrGlnGluLeuArgAla	80
Db	241	CTGTCTGACCAAGTGCAGGAGAAGCGCTCACCTCCACAGTCACCAGGAATCGAAGGGC	300
OY	81	LeuMetAspGLIThrMetLysGluLeuLysAlaTYrLysSeGIuLeuGluGluGluLeu	100
Db	301	CGATGTGACAGACCATGAAAGAGTTGAAGGCCCTACAATAATCGGAATCGAGGAACAAC	360
OY	101	ThrProValAlaGluGluIuThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCCGGTGGCGGAGAGACCGCGGCACGCTCTCCAAAGACTCAGCCAGCGCGCAGGGC	420
OY	121	ArgLeuGluValAspMetLysAspValLysGlyArgLeuValGluTYrArgGlyGluVal	140
Db	421	CGGCTGGGGGGCGACATGGAGACGTGTGGCGCCCTGTGTACAGTACCGCGGCAAGTG	480
OY	141	GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCCATGCTCGGCCAGAGCACCGAGAGAGCTGGGGGTGCCCTCGCGCTCCACCTGGCGC	540
OY	161	LysLeuArgLysArgArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTYr	180
Db	541	AAGCTGCGTAACCGGCTCTCCCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGC	600
OY	181	GlnAlaGlyAlaArgGluGluGluValArgLysArgLysSerAlaIleArgGluArgLeuGly	200
Db	601	CAGGCCGGGGCCCCCGGAGGGGCCCCAGCGCGGCTCAGACGCCATCCGCGAGCGCTGGGG	660
OY	201	ProLeuValGluGluGlnLysArgValArgValArgAlaIleArgValGlySerLeuAlaGlyInPro	220
Db	661	CCCCTGTGTGAGAACAGAGCGCGCGCTGGCGGGCGCCACTGTGTGGCTCTCCGTGGCGGACGCG	720
OY	221	LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluGluMetGly	240
Db	721	CTACAGGAGAGCGGGCCCCAGGCTGTGGGGCGACGCGCTGCCCGGATGAGAGATTGGCGC	780
OY	241	SerArgTYrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys	260
Db	781	AGCCGAGACCCCGGACCGCTGTGACAGAGGTGAAGAGACAGTGGCGAGGTGGCGCCCAAG	840
OY	261	LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleArgAla 277	
Db	841	CTGAGGAGAGAGAGGCCAGACAGATACGCTGTGACGCGCGAGCCCTTCACAGGCC 891	
RESULT 5			
HOMAPOE3			
LOCUS		1156 bp	mRNA linear PRI 24-NOV-2000
DEFINITION	Homo sapiens preapoliipoprotein E (APOE) mRNA, complete cds.		
ACCESSION	K00396		
VERSION	K00396.1 GI:178850		
KEYWORDS	apolipoprotein; apoliipoprotein E; lipoprotein; polymorphism; very low density lipoprotein.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 355 to 1156)		
JOURNAL	Breslow,J.L., McPherson,J., Nussbaum,A.L., Williams,H.W.,		
MEDLINE	Lojquist-Kahl,F., Karachanis,S.K. and Zannis,V.I.		
PUBMED	Identification and DNA sequence of a human apoliipoprotein E cDNA clone.		
	J. Biol. Chem. 257 (24), 14639-14641 (1982)		
	83082756		
	6897404		
	2 (bases 250 to 777)		
	Wallis,S.C., Roane,S., Gill,L., Markham,A., Edge,M., Woods,D.,		
	Williamson,R. and Humphries,S.		
	The isolation of cDNA clones for human apoliipoprotein E and the		
	detection of apoE RNA in hepatic and extra-hepatic tissues		
	EMBO J 2 (12), 2369-2373 (1983)		
	84131952		
	6199196		

REFERENCE AUTHORS	3 (bases 1 to 1156) Zannis-V.I., McPherson,J., Goldberger,G., Karathanasis,S.K. and Breslow,J.L.	CDS	61..1014 /gene="APOE" /codon_start=1 /product="preapolipoprotein E" /protein_id="AA859546.1" /db_xref="GI:178851" /db_xref="GDB:G00-119-691" /translation="MKVLMALVTFPLAGCOAKVEQAVETEPPELPKQTEWMSGORM ELAGRPMDYLRWVOTLSEVOEELSSQVQELRALMDETRKELKAYKSELOLP VAETPRARLSKELQAAQARIGADMDYCGRLVQYRGVQAMIGOSTEELRYRLASHLR KLKRLRLDADDLQRLAVYQAGARREGARERLSLIRELPGLVEQGRRAATVTSLSAG QPLDERAOWAEERLRAREWEGSRTPRDLDEKQVAEVRKLEEQAOQIRLQLEAFO ARLSWEEPLVEDMQRWAGLVEKVOAAVGTSAAPVPSDNH" 61..114 /gene="APOE" /note="apolipoprotein E" 115..1011 /gene="APOE" /product="apolipoprotein E mature peptide" 267 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]" /citation=[2] /replace="c" 279 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]" /citation=[4] /replace="a" 288 /gene="APOE" /note="g in [J. Biol. Chem. 258, 11422-11422 (1983)], [3]" /citation=[4] /replace="a" 369 /gene="APOE" /note="in variant" /citation=[4] /replace="a" 409 /gene="APOE" /note="in variant" /citation=[4] /replace="a" 448 /gene="APOE" /note="in type III hyperlipoproteinemia" /replace="c" 538 /gene="APOE" /note="in type III hyperlipoproteinemia" /replace="c" 568 /gene="APOE" /note="in variant" /citation=[4] /replace="c" 586 /gene="APOE" /note="in epsilon-2 allele" /citation=[3] /citation=[5] /replace="c" 702 /gene="APOE" /note="c in [J. Biol. Chem. 258, 11422-11422 (1983)], [3], [4]" /citation=[2] /replace="g" 783 /gene="APOE" /note="in variant"
TITLE	Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing of a new variant	sig_peptide	
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	J. Biol. Chem. 259 (10), 6498-6504 (1984) 6327682 5 (bases 577 to 624) Gill,L.L., Peoples,O.P., Pearson,D.H., Robertson,F.W., Humphries,S.E., Cumming,A.M. and Hardman,N. Isolation and characterisation of a variant allele of the gene for human apolipoprotein E	mat_peptide	
TITLE	Human apolipoprotein E	variation	
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Biochem. Biophys. Res. Commun. 130 (3), 1261-1266 (1985) 85279526 2992507 6 (sites) Rall,S.C. Jr., Newhouse,Y.M., Clarke,H.R., Weisgraber,K.H., McCarthy,B.J., Mahley,R.W. and Bersot,T.P. Type III hyperlipoproteinemia associated with apolipoprotein E phenotype E3/3. Structure and genetics of an apolipoprotein E3 variant	variation	
JOURNAL MEDLINE PUBMED COMMENT	J. Clin. Invest. 83 (4), 1095-1101 (1989) 89198059 2539388 [1] corrected in [J. Biol. Chem. 258, 11422-11422 (1983)]. [J. Biol. Chem. 258, 11422-11422 (1983)] correction of [1]. [4] epsilon-3 and variant. [5] epsilon-2 allele. [6] sites: mutations resulting in type III hyperlipoproteinemia. Apo E is a component of normal human very low density lipoprotein. There are six human apo E phenotypes known to result from a single structural gene, three of the common alleles being epsilon-4, epsilon-3 and epsilon-2. This sequence appears to be of the epsilon-3 allele. [1] argues that the apo E polymorphism involves mutations in the structural coding region; for example the epsilon-2 phenotype which is characterized by hyperlipoproteinemia is thought to result from a c to t change (arg to cys) at base 586 below [3]. [5]. The sequence shown is 578 homologous with human apo A-1 and 818 homologous with rat apo E. For the epsilon-4 sequence, see the separate entry. [J. Biol. Chem. 258, 11422-11422 (1983)] is too extensively revised relative to [1] to record the revisions in the FEATURES table. The sequence below is that of [J. Biol. Chem. 258, 11422-11422 (1983)] and [3]. Apo E is located on chromosome 19 --Jackson, Bruns and Breslow, PNAS USA 81, 2945-2949 (1984) -- and is believed to be linked to the Apo C-II gene (see separate entry). The two mutations causing type III hyperlipoproteinemia (apolipoprotein E phenotype E3/3) produces substitutions of Arg for Cys at amino acid 112 and Cys for Arg at amino acid position 142. Draft entry and printed copy of sequence for [1] kindly provided by L.L. Gill, University of Aberdeen. Complete source information: Human liver [1] [J. Biol. Chem. 258, 11422-11422 (1983)], [2], [3], [4] and blood [5], cDNA to mRNA. Location/Qualifiers 1..1156 /organism="Homo sapiens" /db_xref="taxon:9606" /map="19q13.2" /tissue_type="liver and blood" 1..1156 /gene="APOE"	variation	
FEATURES source	gene	variation	


```
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9
```

```
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0
```

US-09-827-854-15_COPY_1_277 (1-277) x BC003557 (1-1186)

QY	1	MethylsValLeuTPRPaIaAlaLeuLeuValThrpheLeuAlaIaCysGlnAlaVal	20
Db	46	ATGAAGGTTCGTGGCTCGCTTCGTGTCACTTCCTGCAGAGATGCCAGCCAAAGTG	105
QY	21	GlucInAlaValGIuThrGIuProGIuLeuArgGlnInThrGIuTrpGlnSer	40
Db	106	GAGCAAGCGGTGGAGACAGAGCCCGACCCGACCTGCGCCAGCAGCCAGTGGCAAGC	165
QY	41	GlyGlnArgTPRGIuLeuAlaLeuGlyAArgpHeTrpAspTyrLeuArgTPRValGlnThr	60
Db	166	GGCCAGCGCTGGAACTGGCACTGGGTGGCTTTTGGGATTAACCTGGCCCTGGGTCAACA	225
QY	61	LeuSerGIuGlnValGlnGlnGlnLeuLeuLeuSerSerGlnValThrpGlnGluLeuArgAla	80
Db	226	CTGTCTGAGCAGGTGCAGAGAGACCTGTGTACGTCCACAGTCACCCAGGAAGTGAAGGCG	285
QY	81	LeuMetAspGIuThrMetIysGlnLeuLysAlaTyrIysSerGIuLeuGlnGlnLeu	100
Db	286	CTGTGTGACCGACACATGAAGATTGAAGCCCTCAAAATCGAATCGAGAGAACACTG	345
QY	101	ThrpValAlaGlnGlnGlnThrArgAlaArgLeuSerIysGluLeuGlnAlaGlnAla	120
Db	346	ACCCCGGTGGCGGAGAGACCGCGGCACGGCTGTCCAAGGAGCTGCAGAGCGCGCAGGCC	405
QY	121	ArgLeuGlnAlaAspMetGluAspValIcysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	406	CGCGTGGCGCGGACATGAGAGACGTGTGGCGCGCGCTGTGGCATACCGCGCGGAGGTG	465
QY	141	GlnAlaMetLeuGlnGlnSerThpGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
Db	466	CAGCGCATGTCTCGGCCAGAGACCCGAGAGACTGTGGGTGGCGCTCCCTCCCACTTCGCG	525
QY	161	LysLeuArgLysArgGluLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	526	AAGCTGGGTAAAGCGGCTCTCTCCGGATGCGAGTACCTGTCAAGAGCCCTGGCAGTGTAC	585
QY	181	GlnAlaGlyAlaArgGlnGlnGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Db	586	CAGCGCGGGCGCCGAGAGGGCGCGAGCGGGCGCTCAAGCCACCGCAGAGCGCTGGGG	645
QY	201	ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro	220
Db	646	CCCCCTGGTGAACAGGGCGCGGTGGCGGGCGCCACTGTGGGTCTCCCTGCGCGCCACCG	705
QY	221	LeuGlnGlnArgAlaGlnAlaATPRGlyGluArgLeuArgAlaArgMetGlnGlnMetGly	240
Db	706	CTACAGAGAGCGGCCACAGGCTGGGGCGAGGGCTGTGGCGCGGAGATGGAGGAATGGGC	765
QY	241	SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys	260
Db	766	AGCGGACCGCGCAGCCGCTGGAAGAGGTGAAGAGCAGAGGGGGAGGAGTGGCGCCAG	825
QY	261	LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla	277
Db	826	CTTGAGGAGAGAGCGCCAGCAGATACGCTCTCAGGCGCAGGCTTTCACAGGCC	876

REFERENCE
AUTHORS
TITLE
JOURNAL

I (bases 1 to 1110)
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
MORIMOTO, H. and TERANISHI, Y.
METHOD FOR PRODUCING AFOLIPROTEIN
Patent: JP 1994315392-A 1 15-NOV-1994;

Source	Location/Qualifiers	Year of Publication	Page Count	Notes
1	110			

BASE COUNT	198 a	355 c	414 g	143 t
ORIGIN				

Alignment Scores:	
Pred. No.:	5,58e-72
Score:	1.677,00
Percent Similarity:	99.64%
Best Local Similarity:	99.64%
Query Match:	99.43%
DB:	6
Gaps:	0
Matches:	111
Mismatches:	276
Conservative:	0
Indels:	1
Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x E08423 (1-1110)

QY	1	MelLysValLeuThrPalaIalaLeuValThrPheLeuNlaGlyCysGlnAlaLysVal	20
Db	15	ATGAAGGTCTGTGGGTGGCTGGTGTGCATATCTGTGCAGGATGCCAGGCCAAGTTG	74
QY	21	GIuGlnAlaValGIuThrGIuProGIuPheGIuLeuArgGlnGlnThrGIuTrpGlnSer	40
Db	75	GAGCAACGGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAAGC	130
QY	41	GIyGlnArgTrpGIuLeuAlaIalaLeuGlyIaArgPheTrpAspIlyLeuArgTrpValGlnThr	60
Db	135	GGCCAGCGCTGGGAACTGGCACTGGGTGCTTTGGATTAACCTGCCTGGGTGCACACA	194
QY	61	LeuSerGIuGlnValGIuGlnGIuLeuLeuSerSerGlnValThrGlnGIuLeuArgAla	80
Db	195	CTGTCTAGCAGAGTGCAGAGAGAGACTGCTCAGCTCCAGGTACCCAGAACTAGAGGGC	254
QY	81	LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGlnLeu	100
Db	255	CTGTATGACACAGACCAAGAAGATTTGAAGCCCTACAAATTCGGAACCGAGGAACAACCTG	314
QY	101	ThrProValIaGlnGIuThrArgAlaArgLeuSerLysGIuLeuGlnAlaIaGlnAla	120
Db	315	ACCCGGTGGCGGAGAGACGGCGGCACGGCTGTCCAAGAGCTGCAGGGCGCCAGGCC	374
QY	121	ArgLeuGlnLysAlaAspMetGluAspValCysGlyIaArgLeuValGlnTyrArgGlyGluVal	140
Db	375	CGCGTGGGCCCGGCAAGGAGAGAGTGTGGCGCGCTGTGTGAGATACCCGGCGAGGTG	434
QY	141	GlnAlaMetLeuGlnSerThrGlnGIuLeuArgValArgLeuAlaSerHisLeuArg	160
Db	435	CAGGCACTGCTGGCGACAGACACCGAGAGACTGGGGTGGCGCTGCCCTCCACACTGGC	494
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspSerLeuGlnLysArgLeuAlaValTyr	180

Accession	Protein	Length
D8	AAAGCTGCGTAAAGCGGCTCTCTCCGCGATGCGCATGACCTGCGAAGCGCCTGGCAATGTAC	554
QY	GI1ALAGLIALAARGGLUGLIALAGLUAARGGLYLEU5ERALLAEARGLUARGLEUGLY	200
D8	CAGCGCGGGGCCCCGGAGGGCGCGCAGCGCGGGCTCAGCGCGCATCGCGAGAGCGCTGGGG	614
QY	PROLEUVALGLUGLINGLYIARGYVALARGALAAIATHRVIALGYSERLEUALAGLYSLPRO	220
D8	CCCCGTGTGGAAACAGGGCGCGGTGGGGGCCGCCACTGTGGGCTCCCTCGCGCGCCAGCGG	674
QY	LEUGLINGLUAARGLAGLIALATRPGLYGLUARGLEUARGALARGMETGLUWETGLY	240
D8	CTACAGGAGACGGGCCCAAGGCTGTGGGGCCAGCGCGCTGCCCCGCGGATGGAGAGATGGGC	734
QY	SERATGTHARARGSPARGLEUASPGLUVALYSGGLUGLINVALIALAGLUALARGALALYS	260
D8	AGCGCGACCCCGGACCGGCTGTGACGAGGTGAAGGAGCAGGTGGCCGAGGTGGCGGCCCAAG	794
QY	LEUGLUGLUGLINALGLINGLINTLEARGLEUGLINALAGLUALAPHEGLNALA	277
D8	CTGGGAGGACAGAGGCCACAGATACGCTGTGACGGCGAGGCGCTTCACAGGCC	845

LOCUS	BD004277	1156 bp	DNA	linear	PAT 31-JAN-2002
DEFINITION	Apo E humanized mammal.				
ACCESSION	BD004277				
VERSION	BD004277.1	GI:18632238			
KEYWORDS	JP 2001017028-A/1.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1156)
Fujita,S., Hamanaka,H., Fukui,Y. and yokoyama,M.
Apo E humanized mammal
Patent: JP 2001017028-A 1 23-JAN-2001;
MITSUBISHI CHEMICAL CORP
OS Homo sapiens (human)

PN	JP	2001017028-A/1
PD	23-JAN-2001	
PF	28-APR-2000	JP 2000128919
PR		
PI	SHINOBU FUJITA, HIROKI HAMANAKA, YUKO FUKUI, MINESUKE YOKOYAMA	PC
	A01667/027, A61K45/00, A61P43/28, A61P43/00, C12N5/10, PC	
	C12N15/00, C07K14/775,	
	C12N5/10, C12R1:91), C12N5/00, C12N15/00, (C12N5/00, C12R1:91) CC	

	Location/Qualifiers
Location/Qualifiers	(61) . . (1011) .
Location/Qualifiers	
1 . 1156	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
208 a	367 c 432 g 149 t

Alignment Scores:	
Pred. NO.:	6.65e-72
Score:	1366.00
Percent Similarity:	99.64%
Best Local Similarity:	99.64%
Query Match:	99.42%
DB:	6
Length:	1156
Matches:	276
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-09-827-854-15_COPY_1-277 (1-277) x BD004277 (1-1156)

QY 1 MetIysValIeuTrpAlaIaleuLeuValThrPheLeuAlaCysGlnAlaIysVal 20

Db 61 ATGAGGCTCTGTGGGCTGCGTGGCTGTCACATTCCTGGCAGGATGCCAGGCCAAGTG 120

QY 21 G L G I N A L A V A L G I U T H R G L U P R O G L U P R O G L U E A R G I N G L I N T H R G L U T R P G L I N S E R 40

Db 121 GAGCAGCGGTGGAGACAGACCGGAGCCGAGCTGCGCCAGCAGACCGAGTGGCAGAGC 180

41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60

Db 181 GGCACGGCTGGCACTGGCACTGGGTCCTTTGGGATTACCTGCCGTGGTGCAGACA 240

61 LeuSerGlnInValGlnGlnIuLeuLeuSerSerGlnValThrGlnGlnIuLeuArgAla 80

Db 241 CTGTTGAGCAGGTTGCCTCAGCTCCAGGTCACCCAGAAGAACTGAGGGCG 300

```
81 LeumetSpgluWmMetLySgluleuLysAla";.,ySSerGluLeuGlugluGluInleu 100
```

Db 301 CTGATGCAGACCATGAAAGGAGTTGAAGGCCATACAAATCGGAACATGGAGGAACTG 360

[illegible]

121 ArqLeuGIvAlaSpMetGIuSpValcVsgIvArqLeuValGIIntvArqGIvGIuVal 140

Db 421 CGGCTGGGGCCGGACATGGAGACGTGTGCGGCCCTGTGTACAGTACCGGGCCAGGTG 480

QY 141 GlnAlaMetLeuGlyInserThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160

Db 481 CAGGCCATGCTCGGGCCAGAGCACCGAGAGCTGGGGGTGGGCTCGGCTCCACCTGGCG 540

QY 161 lysleuarglysargleuleuargaspalaaspaspplenglnlysargleualavaltyr 180

Db 541 AAGCTGCGTAAGCGCTCCCTCCGGATGCCGATGACCTGCAGAAGTGGCTGGCAGGTAC 600

QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgGluGly 200

Db 601 CAGGCCGCGGGCCCCGAGGGGCGCCCTCAGGCCATCCGCGAGGCCCTGGGG 660

[illegible]

Cc	3231	<code>TGCGTGGTGAGGAAATGCCTGACGTCGTTGAGCCTCCGTGGCGGCGGCAAGCCG</code>	/200
Dd	661	<code>CCCCGAGGGAACAGGGCCCCTGCGGGGCCACCACGTGTGGCTCCCGTGGCGGCGGCAAGCCG</code>	/200

[illegible]

241 SerArqThrrArqspArqlEuaspGlIValIyscGlInValAlaGlIValAlaArgAlaIys 260

Db 781 AGCCGACCCGGCAGCCCTGACGAGTGAGCAGCAGTGCCGAGCTGCCGCCAAG 840

QY 261 LeuGIuGIuGlnaGIuGlnIleArgLeuGIuGlnaGIuGlnaIlePheGlnaIa 277

Db 841 CTGGAGGACGAGGCCAGCAGATACGCCTGCAGGCGGAGGCTTCCAGGCC 891

RESULT 9
AX333278

LOCUS	AX333278	1157 bp	DNA	linear	PAT 09-JAN-2010
DEFINITION	Sequence 3787 from Patent WO0194629.				

ACCESSION	AX333278
VERSION	AX333278.1
	GI:18123912

KEYWORDS human .

	Homo sapiens
ORGANISM	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Insectivora; Homo

REFERENCE
1
Young D F Augustus M Carter K C Ehner P Endress G
Mammalia; Eumetrida; Filicetes; Cactaceae; Mollusca; nomo.

TITLE Cancer gene determination and therapeutic screening using stromal cells

gene sets
Patent: WO 0194629-A 3787 13-DEC-2001;
JOURNAL

FEATURES

Avalon Pharmaceuticals (US)
Location/Qualifiers

```
source 1. .1157
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
```

BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:

Pred. No.: 7 62e-72
Score: 1365.00
Percent Similarity: 99.28%
Best Local Similarity: 99.28%
Query Match: 99.34%
DB: 6

Length: 1157
Matches: 275
Conservative: 0
Mismatch: 2
Indels: 0
Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AX333278 (1-1157)

QY 1 MettysValleuThrpAlaLeuValThrpheleuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGGCTCGTGTGTCATCTCTGCGAGAGAGCCAGCCCAAGTG 121
QY 21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGCGCGGAGCCGAGCTGCGCAGACAGACGAGTGCGAGAC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GGCGAGCGCTGGAGACTGGACACTGGCTGCTTTGGATTACCTGCGGTGGAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACAGCTGACAGAGAGAGCTGCTCAGCTCCCAAGTCAACCAAGTGAAGGCG 301
QY 81 LeuMetaspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGAGACAGACCATGAAAGAGTTGAAGGCTTCAAAATCGAAGTGGAGACAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 362 ACCCGGAGTGGAGAGAGCGGCGACGCTGTCCAAAGACTCCAGACGCGGAGAGCG 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGCGGCGCTGCTGCTACCTACCGCGGAGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATGCTGCGCAGAGACCGAGAGCTGCGGCTGCGCTCCACCTGCGCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspPheGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAGCGGCTCTCCGCAATCCGATGACTGCAAGACGCGCTGCGACTGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCATCCGCGAGGCGCTGCGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGCGCTGCGGCGCGGATGAGAGATGCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaThrPcLysGluArgLeuArgAlaArgMetGluGlnMetGly 240
DB 722 CTACAGAGAGCGGCGCGAGCGCTGCGGCGCGCTGCGGCGCGGATGAGAGATGCGC 781
QY 241 SerArgTrpArgAspArgLeuAspGluValLysGluGlnValAlaGlnValArgAlaLys 260
DB 782 ACTGAGAGCGGCGCGAGCGCTGCGGCGCGCTGCGGCGCGGATGAGAGATGCGC 841
QY 261 LeuGluGluGlnAlaGlnGlnThrArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 842 CTGGAGAGAGCGCGCGAGAGTACGCTGCGGCGCGGAGCGCTTCCAGAGCG 892

RESULT 10
AX409597
LOCUS AX409597
DEFINITION Sequence 2244 from Patent W00229103.
1157 bp DNA linear PAT 14-JUN-2002

ACCESSION AX409597
VERSION AX409597.1 GI:21442302
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1
Alvarres,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS
TITLE
JOURNAL
Gene expression profiles in liver cancer
Patent: WO 0229103-A 2244 11-Apr-2002;
GENE LOGIC INC (US)

FEATURES
source
1. 1157
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M12529"

BASE COUNT 212 a 370 c 426 g 149 t
ORIGIN

Alignment Scores:

Pred. No.: 7 62e-72
Score: 1365.00
Percent Similarity: 99.28%
Best Local Similarity: 99.28%
Query Match: 99.34%
DB: 6

Length: 1157
Matches: 275
Conservative: 0
Mismatch: 2
Indels: 0
Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AX409597 (1-1157)

QY 1 MettysValleuThrpAlaLeuValThrpheleuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGGCTCGTGTGTCATCTCTGCGAGAGAGCCAGCCCAAGTG 121
QY 21 GluGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGCGCGGAGCCGAGCTGCGCAGACAGACGAGTGCGAGAC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 182 GGCGAGCGCTGGAGACTGGACACTGGCTGCTTTGGATTACCTGCGGTGGAGACA 241
QY 81 LeuMetaspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGluGlnLeu 100
DB 302 CTGATGAGACAGACCATGAAAGAGTTGAAGGCTTCAAAATCGAAGTGGAGACAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
DB 362 ACCCGGAGTGGAGAGAGCGGCGACGCTGTCCAAAGACTCCCAAGTCAACCAAGTGAAGGCG 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyVal 140
DB 422 CGGCTGGGCGGACATGAGAGAGCTGTGCGGCGCTGCTGCTACCTACCGCGGAGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCGCATGCTGCGCAGAGACCGAGAGCTGCGGCTGCGCTCCACCTGCGCG 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspPheGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAGCGGCTCTCCGCAATCCGATGACTGCAAGACGCGCTGCGACTGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCATCCGCGAGGCGCTGCGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
DB 662 CCCCTGTGTGAACAGGCGCGGCTGCGGCGCGCTGCGGCGCGGATGAGAGATGCGC 721

QY 221 LeuGluGluAArgAlaGlnAlaTrrPglYgluAArgLeuAArgAlaArgMetGluMetGly 240
 DB 722 CTACAGAGAGGGGCCAGGCGCTGGGGGAGACGGCGCTGGCGGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValYgluGluGlnValAlaGluValAlaArgAlaYls 260
 DB 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGACAGGTGGCGAGGTGGCGCCAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 DB 842 CTGAGAGAGAGGCCCAAGCAGATACGCTTCAGCGGAGGCTTCCAGGCC 892
 RESULT 11
 LOCUS 115975 1157 bp DNA linear PAT 03-APR-1996
 DEFINITION Sequence 1 from patent US 5472858.
 ACCESSION 115975
 VERSION 115975.1 GI:1250883
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS Attle,A.D., Gretch,D.G., Sturley,S.L. and Beckage,N.E.
 TITLE Production of recombinant proteins in insect larvae
 JOURNAL Patent: US 5472858-A 1 05-DEC-1995;
 FEATURES
 source 1..1157
 ORIGIN /organism="unknown"
 BASE COUNT 212 a 370 c 426 g 149 t

Alignment Scores:
 Pred. No.: 7,62e-72 Length: 1157
 Score: 1365.00 Matches: 275
 Percent Similarity: 99.288 Conservative: 0
 Best Local Similarity: 99.288 Mismatches: 2
 Query Match: 99.348 Indels: 0
 DB: Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x 115975 (1-1157)

QY 1 MetLysValLeuTrrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 DB 62 ATGAAGTTCTGTGGCTGGTGTGTCATCTCTGCGACAGATGCCAGGCAAGTG 121
 QY 21 GluGlnAlaValGluTrrGluProGluProGluLeuArgGlnGlnThrGluTrrPglIns 40
 DB 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGCGCAGACCGAGTGGCAGAGC 181
 QY 41 GlyGlnArgTrrPgluLeuAlaLeuGlyArgPheTrrPAspTrrLeuArgTrrPValGlnThr 60
 DB 182 GCGCAGCGCGGGAAGTGGACACTGGGTCTTTGGGATTACTGGCGCTGGTCAGACA 241
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 DB 242 CTGTCTGAGAGAGGTGACAGAGAGCTGCTACAGTCCCAAGTACCCAGAACTGAGGGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrrLysSerGluLeuGluGluGlnLeu 100
 DB 302 CTGATGAGACGACCATGAAGGAGTTGAAGGCTTACAAATCGGAAGTGGAGGAACTG 361
 QY 101 ThrProValAlaGluGluTrrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 362 ACCCGCGTACCGAGAGACAGCGGCGACGCGTCCAGAGAGTGCAGAGCGGCGAGGCC 421
 QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrrArgGlyGluVal 140
 DB 422 CGGCTGGGGCGGACATGAGGAGCGTGTGGCGCGCTGTGACGTAACGCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160

DB 482 CAGCCATGCTCGGCCAGACACCGAGAGCTCGGGTCCGCTCCCACTCCGCC 541
 QY 161 LysLeuAArgLysArgLeuLeuAArgAspAlaAspAspLeuGlnLysArgLeuAlaTyr 180
 DB 542 AAGCTCGTAAGAGGCGCTCCCGCGATCCCGATGACCTCGACAGAGCGCTGGCAGTAC 601
 QY 181 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 DB 602 CAGCGCGGGCGCCGACAGGCGCGCCGAGCGCGCTCAGGCCCATCCGACAGCGCTGGGG 661
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 662 CCCCTGTGAGACAGGCGCGCGCTGCGCGCCGCTGCTGCTCCCTGGCCGCGCAGCG 721
 QY 221 LeuGluGluAArgAlaGlnAlaTrrPglYgluAArgLeuAArgMetGluMetGly 240
 DB 722 CTACAGAGAGGGGCCAGGCGCTGGGGGAGACGGCGCTGGCGGATGAGAGATGGGC 781
 QY 241 SerArgThrArgAspArgLeuAspGluValYgluGluGlnValAlaGluValAlaArgAlaYls 260
 DB 782 AGTCGAGACCGCGACCGCTGGAGAGGTGAAGAGACAGGTGGCGAGGTGGCGCCAG 841
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 DB 842 CTGAGAGAGAGGCCCAAGCAGATACGCTTCAGCGGAGGCTTCCAGGCC 892
 RESULT 12
 LOCUS HUMAPOE 1157 bp mRNA linear PRI 08-AUG-1995
 DEFINITION Human apolipoprotein E mRNA, complete cds.
 ACCESSION M12529
 VERSION M12529.1 GI:178848
 KEYWORDS
 SOURCE Homo sapiens (clone: PHAE[112,178,813].) male 57-year old liver
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1157)
 AUTHORS McLean,J.W., Elshourbagy,N.A., Chang,D.J., Mahley,R.W. and Taylor,J.M.
 TITLE Human apolipoprotein E mRNA. cDNA cloning and nucleotide sequencing
 JOURNAL J. Biol. Chem. 259 (10), 6498-6504 (1984)
 MEDLINE 84212473
 PUBMED 6327682
 FEATURES
 source 1..1157
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19q13.2"
 /clone="PHAE[112,178,813]."
 /sex="male"
 /tissue_type="liver"
 /dev_stage="57-year old"
 1..1157
 /gene="APOE"
 <1..1157
 /note="apoE mRNA; 600-119-691"
 62..1015
 /gene="APOE"
 /note="precursor"
 /codon_start=1
 /product="apolipoprotein E"
 /protein_id="AAB59518.1"
 /db_xref="GI:178849"
 /db_xref="GDB:600-119-691"
 /translation="MKVLMALLVFLFAGCOAKVEAOAVETREPEELROOTMOSGORM
 ELALGFMDYLRMVOYLSEVOEELLSVOYELALNDETMKELKAVKSELEQLTP
 VAETRRRLSEIOTFAORILGMDVCGRLVYQVGEVQAMIGOSTEELRVRLSHLR
 KLRRLDPPDLLKRLAVYQAGREGNERGISATREHLGPLEVDGRRVRAATVSLAG
 OPLERRAOWGBRILKARMEWGSRTDRDLDEVKEDVAVRAKLEBOAOOIRLOAEAFQ

QY	21	GIUGINAlaValaIGluhrngIuprpgIuuprpgIuAArgInsgInhrngIuprpgInser	40
Db	121	GAGCAAGCGGTGGAGACAGAGCCGGAGCCGAGCTCGCCAGCAGACCGAGTGGCAGAG	180
QY	41	GIyGInArgrTpDluenuAlaIeugIyArPheTrpAsPtyrIeuargrTpaValGInhr	60
Db	181	GGCCACGCGTGGGAACCTGGGACTGGGTCCGCTTTGGGCAITTTACCTGGCGTGGCGAGACA	240
QY	61	LeuSerGIuGInVaIcInGInGluLeuLeuSerSerGIuValThrGInGluLeuAArgAla	80
Db	241	CGTCTGACACAGGTGCAGAGAGAGCTGCATCAGCTCCAGGTCCACCAGAACTGAGAGGGG	300
QY	81	LeuMetAspGIuThrMetIySGIuIeuIySAIaTyIySerGIuIeugIuGInGInIeu	100
Db	301	CTGATGACAGACCCATGAAAGAGTGAAGCCCTTACAAATTCGAACTGAGAGAAACAATG	360
QY	101	ThrProValAlaGIuIuIuThrArGaIaArGrLeuSerIySGIuIeugInAlaIaGIuAla	120
Db	361	ACCCCGGTGGCGAGAGACCGCGGACCGCTGTCCAAAGAGCTGCAGGCGGCGAGGCC	420
QY	121	ArgIeugIyAlaAspMetGIuAspValCySGIyArgIeUValGIuIuThrArGIyGInuVal	140
Db	421	CGGCTGGCGCGGACATGAGAGACGTGGCGGCGCGCTGTGTGAGTACCGCGGCGAGGTG	480
QY	141	GIuAlaMetIeugIyGInSerThrGIuIeUArGValArgIeUAlaSerHisIeuarG	160
Db	481	CAGGCGATCTCTGGCCAGACACCGAGAGACTCGGGGTCCGCTCCCTCCACCTGGCC	540
QY	161	LysIeUArGIyArgrIeUArGrAspAlaAspAspIeUcInLysArgrIeUAlaValTyrr	180
Db	541	AGGCTGCGTAAAGGGGTCTCTCCGCGATGTCGATGACTTCAGAAAGCGCTGGCAGGTAC	600
QY	181	GIuAlaGIyAlaArgIuIuIyAlaGIuArGIyIeUSerAlaIeuarGIuArGrIeugIy	200
Db	601	CAGGCGGGGCGCCGAGAGGCGCGGAGCGCGGTCTCAGCGCATCCGAGGCGGCGTGGG	660
QY	201	ProIeUValGIuGInGIyArGValArgAlaAlaThrValGIySerIeUAlaGIuIuPro	220
Db	661	CCCTGTGTGAACAGAGGCGCGGTGGCGGCGGCTGCTGCTGCTGCTGCGGCGAGCGG	720
QY	221	LeuGIuGIuArGIaGIuAlaTrpGIyGIuArGrIeUArGIaArgMetGIuIuMetGIy	240
Db	721	CTACAGAGACGGGCCCCAGGCGCTGGGGCGAGCGGTCTGCGCGCGATGGAAGAGATGGG	780
QY	241	SerArGrThrArGIAspArGrIeUAspGIuValIySGIuGIuAlaGIuValArGIaIyS	260
Db	781	ACCCGAGACCGCGAGACCGCTTGAGACAGGTGAGAGAGACAGTGGCGAGGTGCGCGCAG	840
QY	261	LeuGIuGIuGIuAlaGIuGIuIleArGrIeUAlaGIuAlaIaPheGInAla	277
Db	841	CTGAGAGAGACAGGCCAGATACGCCCTGCGAGCGGAGGCTTCACAGGCC	891
RESULT 14			
MEAPOE			
LOCUS	MEAPOE	1178 bp	mRNA
DEFINITION	Monkey mRNA for apolipoprotein E.		linear
ACCESSION	X13887		
VERSION	X13887.1	GI:38054	
KEYWORDS	apolipoprotein; apolipoprotein E.		
SOURCE	Macaca fascicularis.		
ORGANISM	Macaca fascicularis		
	Eumalvota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;		
	Cercopithecoidea; Macaca.		
	1 (bases 1 to 1178)		
REFERENCE	Marotti,K.R.		
AUTHORS	Direct Submission		
TITLE	Submitted (16-DEC-1988) Marotti K.R., The Upjohn Co, 7242 209 713,		
JOURNAL	301 Henrietta Street, Kalamazoo, MI 49008		
REFERENCE	2 (bases 1 to 1178)		
AUTHORS	Marotti,K.R., Whitted,B.E., Castle,C.K., Polites,H.G. and		
	Meilhoir,G.W.		
TITLE	Nucleotide sequence of the cynomolgus monkey apolipoprotein E cDNA		

Journal	Nucleic Acids Res.	17 (4), 1778 (1989)
Medicine	89160349	
PubMed	2922300	
Comment	Data kindly reviewed (20-Mar-1989) by Marotti K.R.	
Features	Location/Qualifiers	
Source	1..1178	
	/organism="Macaca fascicularis"	
	/db_xref="taxon:9541"	
	/clone="1"	
	/tissue_type="liver"	
	/clone.lib="lambda gt11"	
	83..1036	
CDS	/codon_start=1	
	/product="proapolipoprotein E (AA -18 to 299)"	
	/protein_id="CAA32092.1"	
	/db_xref="GI:38055"	
	/db_xref="SWISS-PROT:P10517"	
	/translation="MKVLMALIVTFPLRQACQAKVQDPPEPEPEPELRLQAAEGSGQPK ELALGRFDVILRWYQTLSEYQBELSQYQELTTLMDFTMKELKATKSLSEQLS VAEETRLRLSKELQAAQRLGADMDVNSRLVQYRSEVQAALGOSTEELRLRLSLRL KLRLRLDADDLQRLAVYQAGREGNERVSAIREDLGLVDSGRVRAFTVSLAS QPLERRAOLGERLRLRAMEEGRSTRDLDEKQVAEVRALKEEQAOQISLAERAC ARKLSWPEPLVEDMQRMAGLVERVQAAYGASTAPVPIDN"	
	83..136	
sig_peptide	137..1033	
mat_peptide	/product="apolipoprotein E (AA 1 - 299)"	
	1154..1159	
misc_feature	/note="polya signal"	
BASE COUNT	217 a 373 c 426 g 162 t .	
ORIGIN		
Alignment Scores:		
Pred. No.:	2,12e-65	Length: 1178
Score:	1255.00	Matches: 259
Percent Similarity:	94.22%	Conservative: 2
Best Local Similarity:	93.50%	Mismatches: 16
Query Match:	91.34%	Indels: 0
DB:	9	Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x MFAOE (1-1178)		
QY	1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
Db	83 ATGAAGGTTCTGTGGGCTGGCTGCTGTGGCAATTCCTGGCAGATGCGCAGGCCAAAGTG	142
QY	21 GlnGlnAlaValAlaGlnThrcLupProGlnPocGlnLeuArgGlnGlnThrcLupTrpIns	40
Db	143 GAGCAACCGGTGGACCCAGACGAGACGCCAGCTTCGCCACGACGCTGAGGGCCAGACC	202
QY	41 GlyIAsArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
Db	203 GGGCAGCCCTGGGAGACTGGCACTGGTGCTTTGGATTACCTGCGTGGGTGCAGACA	262
QY	61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrcLingLuleuArgAla	80
Db	263 CTGTCTGACAGGTGTACAGAGAGAGCTGCTCAGCCGCCAGGTCACCCAGAGACTGACGAGC	322
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnLeu	100
Db	323 CTGATGGACGAGACCATGAGAGGTTGAAGGCTTCAACAATCGGAATCGAGAGACAGCTG	382
QY	101 ThrProValAlaGlnGlnGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	383 AGCCCGGGTGGCGAGAGACACGGGGCAGCGCTGTCCAGAGACGTGACAGGGCCAGGCC	442
QY	121 ArgLeuGlyAlaAspMetCylAspValCysGlyArgLeuValGlnTrpArgGlyVal	140
Db	443 GGGCTGGGTGGCGAGCATGAGAGACTGCGCAGACGCCCTGTGTGACGTACCGCAGCGAGTG	502
QY	141 GlnAlaMetLeuGlyGlnSerThrcLingLuleuArgValArgLeuAlaSerHisLeuArg	160
Db	503 CAGGCATGCTGGGCGACAGATPACGAGGAGCTGGGGCGCGCTGCGCTCCACACTGCGC	562

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:08:17 ; Search time 156.176 Seconds
(without alignments)
3994.237 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
1374
Sequence: 1 MKVLMALVLTFLAGCOAKV.....RAKLEQQAQQRIGQAEAFQA 277

Scoring table:
BLSDUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+._p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09827854/funat_11032003_101609_27476/app_query.fasta_1.3576
-DB=N.Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blscomb2 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827854.ecgn1.1_1201_etrnat_11032003_101609_27476 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MMAP -LARGEQUERY -NES-SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_101002.*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1374	100.0	954	AAD26035	Human apolipoprotein E
2	1374	100.0	1110	AAH60409	Human apolipoprotein E
3	1374	100.0	1147	ABA83113	Human apolipoprotein E
4	1374	100.0	1156	AAH64315	Human apolipoprotein E
5	1374	100.0	1156	AAH22048	Human apolipoprotein E
6	1370	99.7	1156	AAH22052	Human apolipoprotein E
7	1366	99.4	1156	AAH64314	Human apolipoprotein E
8	1366	99.4	1156	AAH22049	Human apolipoprotein E
9	1366	99.4	1156	AAH22051	Human apolipoprotein E
10	1365	99.3	1157	AAH6957	Human apolipoprotein E
11	1365	99.3	1157	ABN95746	Human apolipoprotein E
12	1365	99.3	1157	ABK64514	Human apolipoprotein E
13	1365	99.3	1157	ABL65450	Human apolipoprotein E
14	1364	99.3	1110	AAH50450	Human apolipoprotein E
15	1362	99.1	1156	AAH64316	Human apolipoprotein E
16	1362	99.1	1156	AAH22047	Human apolipoprotein E
17	1359	98.9	1156	AAH22050	Human apolipoprotein E
18	1322	96.2	1279	AAH22437	Human apolipoprotein E
19	1251.5	91.1	1107	AAH75756	Human apolipoprotein E
20	1185.5	86.3	9360	AAH31915	Human apolipoprotein E
21	1185.5	86.3	10716	AAH26034	Human apolipoprotein E
22	1180	85.9	3805	AAH209524	Human apolipoprotein E
23	1180	85.9	3805	AAH209526	Human apolipoprotein E
24	1158.5	84.3	10716	AAH26108	Human apolipoprotein E
25	1004.5	73.1	965	AAH32081	Human apolipoprotein E
26	1004.5	73.1	5617	AAH32077	Human apolipoprotein E
27	1004.5	73.1	6026	AAH32075	Human apolipoprotein E
28	1002.5	73.0	1126	AAH29159	Human apolipoprotein E
29	984	71.6	660	AAH69792	Human apolipoprotein E
30	919	66.9	936	AAH69101	Human apolipoprotein E
31	919	66.9	936	AAH8070	Human apolipoprotein E
32	881	64.1	597	AAH18068	Human apolipoprotein E
33	872	63.5	597	AAH69099	Human apolipoprotein E
34	823.5	59.9	1381	AAH22673	Human apolipoprotein E
35	678	49.3	600	AAH89595	Human apolipoprotein E
36	656.5	47.8	786	AAH18114	Human apolipoprotein E
37	651	47.4	407	ABK34238	Human apolipoprotein E
38	625	45.5	478	ABN95994	Human apolipoprotein E
39	625	45.5	478	ABH2679	Human apolipoprotein E
40	625	45.5	478	ABH2679	Human apolipoprotein E
41	594	43.2	499	AAH63049	Human apolipoprotein E
42	514	37.4	330	AAH11980	Human apolipoprotein E
43	504	36.7	405	AAH20139	Human apolipoprotein E
44	485	35.3	345	AAH8479	Human apolipoprotein E
45	419	30.5	260	AAH40342	Human apolipoprotein E

ALIGNMENTS

RESULT 1
AAD26035
ID AAD26035 standard; cDNA: 954 BP.

AC AAD26035;
XX 26-MAR-2002 (first entry)
XX Human apolipoprotein E (APOE) cDNA.

DE Human; anti-lipemic; neuroprotective; nootropic; genetic variant; APOE;
XX apolipoprotein E; haplotyping; familial dysbetalipoproteinemia; therapy;
KW genotyping; type III hyperlipoproteinemia; Alzheimer's disease; SNP;
KW atherosclerosis; single nucleotide polymorphism; Chromosome 19q13.2; ss.
XX Homo sapiens.
OS Location/Qualifiers
FH key

FT	CDS	1..954
FT	/*tag= a	
FT	/product= "Human APOE protein"	
FT	replace (13, C)	
FT	/*tag= a	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (31, G)	
FT	/*tag= b	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (364, A)	
FT	/*tag= c	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (388, C)	
FT	/*tag= d	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (498, G)	
FT	/*tag= e	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (526, T)	
FT	/*tag= f	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
FT	replace (622, A)	
FT	/*tag= g	
FT	/standard_name=	"Single nucleotide polymorphism (SNP) "
PN	WO200179234-A2.	
PD	25-OCM-2001.	
XX		
XX	16-APR-2001; 2001WO-US12303.	
PF		
XX	14-APR-2000; 2000US-197188P.	
PR		
XX	(GENA-) GENMASSANCE PHARM INC.	
PA		
XX		
XX	Choi JY, Kilem SE, Koshy B, Lee HH;	
PI		
DR	WPI: 2002-075064/10.	
XX	P-PSDB; AAE15158.	
XX		
PT	Genotyping human apolipoprotein gene of individual for determining	
PT	haplotype of individual, involves determining identity of nucleotide	
PT	pair at specific polymorphic sites for two copies of gene -	
PS	Claim 26; Fig 2; 78pp: English.	
XX		
CC	The patent discloses novel genetic variants of human apolipoprotein	
CC	E (APOE) gene. The invention also relates to compositions and methods	
CC	for haplotyping and/or genotyping the APOE gene. The haplotyping	
CC	methods of the invention are useful for improving the efficacy and	
CC	reliability of several steps in the discovery and development of	
CC	drugs for treating diseases associated with APOE activity, e.g.	
CC	familial dysbetalipoproteinemia, type III hyperlipoproteinaemia,	
CC	atherosclerosis, and Alzheimer's disease. They are useful to validate	
CC	APOE as a candidate agent for treating a specific condition or disease	
CC	predicted to be associated with APOE activity and in the design of	
CC	clinical trials of candidate drugs for treating a specific condition	
CC	or disease predicted to be associated with APOE activity. Genotyping	
CC	or haplotyping methods are useful to screen for compounds targeting	
CC	APOE to treat a specific condition or disease associated with APOE	
CC	activity. The present sequence is a cDNA encoding human APOE protein.	
CC	APOE gene is located on chromosome 19q13.2.	
XX		
SO	Sequence 954 BP; 168 A; 290 C; 377 G; 119 T; 0 other;	
Alignment Scores:		
Pred. No.:	1.84e-102	Length: 954
Score:	1374..00	Matches: 277
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	24	Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AAD26035 (1-954)
QY 1 MetLysValLeuTPtAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIysVal 20
Db 1 ATGAAGGTTCTGTGGGCTGGCTTCTGTCACTTCTCTGGCAGGATCCAGGCCAAGGTG 60
QY 21 GlnGlnAlaValGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnIleThrPheLeuSer 40
Db 61 GAGCAGACGGCTGGAGACAGACCCGAGCCGAGCTGGCCAGCAGACCGAGTGGCAGAGC 120
QY 41 GlyIleArgTrpGluLeuAlaLeuGlyIleArgPheThrPaspTrpLeuArgTrpValGlnThr 60
Db 121 GGCCAGCGCTGGAACTGGCACTGGTGGCTTTGGGATTACCTGGCTGGTGGTGGAGACA 180
QY 61 LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 181 CTGTCTGGACAGGTGCAGAGAGAGACTGCTCAAGCTCCAGAGTCAACCCAGGAACAGAGGGCG 240
QY 81 LeuMetAspGluThrMetLysGlnLeuLysAlaTrpLysSerGlnLeuGlnGlnGlnLeu 100
Db 241 CTGATGGACGAGACCATGTAAGAGATTGAAGGCTTCAAAATCGGAACCTGGAGGAACAATG 300
QY 101 ThrProValAlaGlnGlnThrThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
Db 301 ACCCGCGTGGGGAGAGAGACCGGGGCAAGGCTGTCCAAGAGGTGCAGCGGGGCGAGGCC 360
QY 121 ArgLeuGlnAlaAspMetGlnAspValCysGlyIleValGlnThrArgGlyGlnVal 140
Db 361 CGGCTGGGCGGCGACATGGAAGAGCTGTGGGGCCCTGTGTGCAGTACCCTGGCGAGGTG 420
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 421 CAGGCCATGCTCGGCCAGAGACCCAGAGAGCTGGGGTGGCGCTCGCTCCACCTGGCC 480
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaIaIyTr 180
Db 481 AAGCTGGCTAAGCGGCTCTCTCCGGATGCCGATGACTCGCAAGAACCCCTGGCAGTGTAC 540
QY 181 GlnAlaGlnAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
Db 541 CAGGCGGGGGCGCGCAGAGGGCGCCGAGCGCGCTCAGCCGCAATCCGCGAGCCCTGGGG 600
QY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValIleLysSerLeuAlaGlyGlnPro 220
Db 601 CCCCTGGTGGAAACAGGGCCCGGTGGGGCCGCCACTGTGGGGTCCCTGGCGGGCCAGCG 660
QY 221 LeuGlnGlnArgAlaGlnAlaIleTrpGlyGlyIleArgLeuArgAlaIleArgMetGlnGlnMetGly 240
Db 661 CTACAGAGAGCGGGGCCAGGCTGTGGGGCGAGCGCTGGCGCCGCGGATGAGAGAGATGGCC 720
QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
Db 721 AGCGGAGACCCCGCAGACCGCTGTGACAGAGTGAAGAGAGCAGGTGGCGGAGTGGCGCCCAAG 780
QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
Db 781 CTGGAGGAGCAGAGCCCGACAGATACGCTCAGGCGCAGAGCCCTTCCAGAGCC 831
RESULT 2
AAAN60409
ID AAAN60409 standard; DNA: 1110 BP.
XX AAAN60409;
XX 01-JAN-1980 (first entry)
XX Human apolipoprotein-E.
XX Human apolipoprotein-E.
XX Apolipoprotein-E; hyperlipidemia; arteriosclerosis; ss.
XX Homo sapiens.
XX Key Location/Qualifiers

FT	CDS	15..968	
FT		/*tag= a	
XX			
XX			
XX	AAU8547513-A.		
XX			
PD	24-APR-1986.		
XX			
XX			
PF	17-SEP-1985;	85AU-0047513.	
XX			
XX			
PR	11-JUN-1985;	85JP-0126989.	
XX			
XX			
PA	(MITU) MITSUBISHI CHEM IND KK.		
DR	WPI: 1986-150217/24.		
DR	P-PSDB: AAP60507.		
XX			
XX			
PT	New DNA sequence coding for human apolipoprotein-E - and		
PT	expression vectors and transformed cells contg. It		
XX			
PS	Disclosure: Fig 2; 45pp; English.		
XX			
CC	The encoded protein is used to treat subjects who are deficient in		
CC	apolipoprotein-E (or who produce abnormal forms of this molecule)		
CC	and therefore are likely to suffer from hyperlipidemia, resulting in		
CC	arteriosclerosis. It can also be used to raise antisera for		
CC	detecting the protein deficiency or production of abnormal forms.		
XX			
XX			
SO	Sequence 1110 BP; 198 A; 354 C; 415 G; 143 T; 0 other:		
	Alignment Scores:		
	Pred. No.:	2.19e-102	Length: 1110
	Score:	1374.00	Matches: 277
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	100.00%	Indels: 0
	DB:	7	Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x AAN60409 (1-1110)			
OY	1 MetlyValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20		
Db	15 ATGAAGGTTCTGTGGCGCTGCTGCTGTCACATTCTGGCAGAGATGCCCAAGGTG 74		
OY	21 GluGlnAlaValAlaGluThrGluProGluProGluLeuArgGlnGlnThrGlnPhe 40		
Db	75 GAGCAAGGCGTGGAGACAGACCGGAGCGCTGCCCGCCAGCAGACCGGAGTGGCAGAC 134		
OY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60		
Db	135 GGCCAGCGCTGGAACTGGCACTGGCTGCTTTGGGATATACCTGCGCTGGGTGCAGACA 194		
OY	61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80		
Db	195 CTGCTGTAGCAGGTGCAGAGAGAGCTGCTCCAGGTCCACCGACGAACTGAGGGCG 254		
OY	81 LeuMetAspGluThrMetLeuGluLeuLeuValTrpLeuSerGluLeuGluGlnLeu 100		
Db	255 CTGATGGAGCAGACCATTAAGAGCTGAAGCCCTACAAATGCGAACTGGAGAACACTG 314		
OY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120		
Db	315 ACCCGCGTGGCGAGAGACCGCGGCGCTGTCCAAAGAGACGACAGCGCGCAGGCC 374		
OY	121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140		
Db	375 CGGCTGGGCGCCGACATGAGAGAGCTGTGCGCGCTGTGTGTCAGTACGGGCGAGGTG 434		
OY	141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValAlaGlnAlaSerHisLeuArg 160		
Db	435 CAGGCGATGCTCGGCGACAGCAGCAGAGAGCTGGGGTGGCCCTCCCTCCACTGGGC 494		
OY	161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180		
Db	495 AAGCTGCGTAAGCGGCTCTCCCGCATGCCATGACCTGCAGAAAGCCCTGGCACTGTAC 554		
OY	181 GlnAlaGlyAlaArgGluGlyValAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200		
Db	555 CAGGCGGGGCGCCCGGAGGGCGCCGAGCGCGGCTCAGCGCCATCCGCGAGCGCTGGGG 614		
OY	201 ProLeuAlaGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220		
Db	615 CCCCTGTGTGAACAGAGCGCGGTGGCGCCGCACTGTGGGTCTCCCTGGCGGCCAGCGC 674		
OY	221 LeuGlnGluArgAlaGlnAlaTrpGlyValArgLeuArgAlaArgMetGluGlnMetLys 240		
Db	675 CTACAGAGAGCGGCGCCAGCGCTGGGGCGAGCGGCTGCCCGCGCATGAGAGATGGGC 734		
OY	241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260		
Db	735 AGCCGAGACCGCGACCGCTGAGCAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCCAAG 794		
OY	261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277		
Db	795 CTGAGAGAGACAGCGCCACAGATACGCTGCAGCGCCGAGCCCTTCCAGGCC 845		
	RESULT 3		
	ABA83113		
	ID ABA83113 standard; DNA; 1147 BP.		
	XX		
	XX ABA83113;		
	XX		
	XX 08-FEB-2002 (first entry)		
	DT		
	XX		
	XX Apolipoprotein E ovarian tumour marker gene, SEQ ID NO:63.		
	DE		
	XX		
	KW Ovarian tumour marker gene; human; overexpression; upregulation;		
	KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;		
	KW identification; serous cystadenoma; borderline serous tumour;		
	KW serous cystadenocarcinoma; mucinous cystadenocarcinoma;		
	KW mucinous cystadenoma; borderline mucinous tumour; endometrioid carcinoma;		
	KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;		
	KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;		
	KW immune response pathway; cell proliferation regulation; protein folding;		
	KW membrane localized; secreted; therapeutic target; cytostatic;		
	KW gene therapy; vaccine; ds.		
	KW		
	XX		
	XX Homo sapiens.		
	XX		
	XX W0200175177-A2.		
	PN		
	XX		
	XX 11-OCT-2001.		
	PD		
	XX		
	XX 03-APR-2001; 2001WO-US10947.		
	PF		
	XX		
	XX 03-APR-2000; 2000US-194336P.		
	PR		
	XX		
	XX (USSH) US DEPT HEALTH & HUMAN SERVICES.		
	PA		
	XX		
	XX Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;		
	PI		
	XX WPI: 2001-626450/72.		
	DR		
	XX P-PSDB: ABB50287.		
	PT		
	XX		
	XX		
	XX		
	PS		
	XX		
	XX Claim 23; Page 105-106; 140pp; English.		
	CC		
	CC The invention relates to methods for diagnosing and prognosing ovarian		
	CC tumours in an individual via the detection and measurement of the		
	CC expression of ovarian tumour marker genes (ABA83081-ABA83122, ABA83180,		
	CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83169, ABA83179,		
	CC ABA83181 and ABA83183). The methods of the invention are useful for		
	CC detecting an ovarian tumour in a patient, for identifying an individual		

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
 DB 61 ATGAGAGTTCTGTGGGTGGTGGTGTGTCACATTCCTGGCAGAGATCCAGGCCAAGGTG 120
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGAGACGAGAGCCAGAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCCACGCGGTGGAGACTGGACAGCGGTGGCTTTGGATTACCTGGCTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGAGGGC 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 DB 301 CTGATGGACAGACCATGAAGGAGTGAAGGCTTACAATACTGGAACCTGGAGAACACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 DB 361 ACCCGGTGGCGAGAGAGACGCGGCGCTGTCCAGAGAGCTGCAGCGCGCGCAGGCC 420
 QY 121 ArgLeuGluValAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyVal 140
 DB 421 CGGCTGGGCGCGACATGAGAGACGTGTGCGCGCTGGTGCAGTACCGCGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 DB 481 CAGGCGATGCTCGGCGAGACACCGAGAGCTCGGGTGGCGCTCGCTCCACCTGGCGC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 DB 541 AACCTCGTAAGCGGTCTCTCCGCGATGCGAGACTTCAGAGAGCGCTGGCAGAGTAC 600
 QY 181 GlnAlaGluValArgGlnGlnGlnArgGlyLeuSerAlaIleArgGluArgGly 200
 DB 601 CAGCGCGGGCGCGCAGGCGCGCGCGCGCTCAGCGCGCTCCGCGAGCGCTCGGG 660
 QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlnPro 220
 DB 661 CCCCTGTGGAAACAGGGCGCGCTGGCGCGCTGTGGGTCCCTGGCGCGCGCAGCGC 720
 QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
 DB 721 CTACAGAGCGGGCGCGCTGGCGCGAGCGCTCGCGCGGAGTGGAGGAGATGGGC 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaIys 260
 DB 781 AGCCGAGACCGCGACCGCTGGACGAGGTGAAGAGACAGCTGGCGGAGCGCCAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 DB 841 CTGAGAGAGCAGGCCAGAGATACCTGCGAGCGCGAGCGCTTCCAGAGCC 891
 RESULT 5
 AAD22048
 ID AAD22048 standard; DNA; 1156 BP.
 XX AAD22048;
 AC AAD22048;
 XX
 DT 12-FEB-2002 (first entry)
 DE Human apolipoprotein E (apoE) isoprotein, apoE3 DNA.
 XX
 KM Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
 XX hypertriglyceridaemia; low density lipoprotein; LDL; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT CDS 61..1014
 FT FT /*tag= a
 FT sig_peptide /product= "Human apoE isoprotein, apoE3"
 FT 61..114
 FT mat_peptide /*tag= b
 FT 115..1011
 FT /*tag= c
 FT /product= "Mature human apoE isoprotein, apoE3"
 PN WO200177136-A1.
 PD 18-OCT-2001.
 PF 06-APR-2001; 2001WO-US11358.
 PR 06-APR-2000; 2000US-0544386.
 PR 04-OCT-2000; 2000US-0679088.
 PR 05-APR-2001; 2001US-0827854.
 XX
 PA (KOSP-) KOS PHARM INC.
 PA (UYBO-) UNIV BOSTON.
 XX
 PI Zannis VI, Kypros KE;
 XX
 DR WPI: 2002-010885/01.
 DR P-PSDB; AME13294.
 PT New apolipoprotein E polypeptide and nucleic acid, useful for lowering
 PT cholesterol, delaying the onset of or treating atherosclerosis in
 PT mammal, without inducing hypertriglyceridaemia -
 PS
 XX
 XX Claim 14; Page 81; 91pp; English.
 CC The present sequence is a human apolipoprotein E (apoE)
 CC isoprotein, apoE3 DNA. The apoE lipoproteins are useful for
 CC lowering cholesterol, delaying the onset of atherosclerosis,
 CC treating or regressing atherosclerosis without inducing
 CC hypertriglyceridaemia, in a mammal lacking an endogenous,
 CC normally functioning apoE gene or low density lipoprotein (LDL)
 CC receptor or is at risk for developing atherosclerosis due to
 CC accumulation of lipoprotein remnants in the bloodstream or having
 CC a defect in remnant removal.
 XX
 SQ Sequence 1156 BP; 208 A; 368 C; 432 G; 148 T; 0 other;
 Alignment Scores:
 Pred. No.: 2.3e-102 Length: 1156
 Score: 1374.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-15_copy_1_277 (1-277) x AAD22048 (1-1156)
 QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaIysVal 20
 DB 61 ATGAGAGTTCTGTGGGTGGTGGTGTGTCACATTCCTGGCAGAGATCCAGGCCAAGGTG 120
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGGCCAGAGACGAGAGCCAGAGAGC 180
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 DB 181 GGCCACGCGGTGGAGACTGGACAGCGGTGGCTTTGGATTACCTGGCTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGGAACTGAGAGGGC 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100

```
Db 301 CTGATGACAGACACCATGAAGGAGTTGAAGCCTACAAATCGGAGACTGGAGCAACTG 360
Qy 101 ThpProValAlaIugIurThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGACGGGGCACGGCTGTCCAAAGAGAGTGCAGAGCGCGGACAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACAGAGGAGCGTGTGGCGCGCTGTGTGCAATCCCGCGGAGAGTG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCACAGACACCGAGAGACTCGGGGTGGCGCTCCGCTCCACCTGCGC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCCTGCGAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGlyGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
Db 601 CAGCGCGGGCGCGCGAGGGCGCGGAGCGCGGCTCAGCCCATCCGCGAGCGCTGGGG 660
Qy 201 ProLeuValGluGlnArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGACAGAGGCGCGCTGTGGCGCGCCACCTGTGGCTCCCTGGCGGCGACCG 720
Qy 221 LeuGlnGluArgAlaGlnAlaIleArgGlyGluArgLeuArgAlaArgMetGluMetGly 240
Db 721 CTNAGAGAGAGGGGCGCGGCTGTGGGGAGCGGCTGTGGCGCGGAGTGCAGAGATGGCG 780
Qy 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCGCGGACCGCGCAGCGCTGTGACAGAGTGAAGAGAGAGTGGCGGAGTGGCGCGCAAG 840
Qy 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaIleGlnAla 277
Db 841 CTGAGAGAGAGCGCCAGCATACGCTGTGAGGCGAGGCGCTTCCAGGCC 891
RESULT 6
AAD22052
ID AAD22052 standard; DNA; 1156 BP.
AC AAD22052;
XX
XX 12-FEB-2002 (first entry)
DE Human apolipoprotein E (apoE) allele, apoE2** DNA.
XX
KW Human: apolipoprotein E; apoE; cholesterol; atherosclerosis;
KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH CDS 61..1014
FH FT /tag= a
FH FT /product= "Human apoE allele, apoE2**"
FH FT sig_peptide 61..114
FH FT /tag= b
FH FT mat_peptide 115..1011
FH FT /tag= c
FH FT /product= "Mature human apoE allele, apoE2**"
XX
PN WO200177136-A1.
PD 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-US11358.
PF
XX
XX 06-APR-2000; 2000US-0544386.
PR
XX 04-OCT-2000; 2000US-0679088.
PR
XX 05-APR-2001; 2001US-0827854.
XX
```

```
PA (KOSP-) KOS PHARM INC.
PA (UYBO-) UNITV BOSTON.
XX
XX Zanniss VI, Kypros KE;
XX
XX WPI: 2002-010885/01.
XX P-PDB: AAEL3298.
XX
XX New apolipoprotein E polypeptide and nucleic acid, useful for lowering
XX cholesterol, delaying the onset of or treating atherosclerosis in
XX mammal, without inducing hypertriglyceridaemia -
XX
XX Claim 14; page 83; 91pp; English.
XX
XX The present sequence is a human apolipoprotein E (apoE)
XX allele, apoE2** DNA. The apoE lipoproteins are useful for
XX lowering cholesterol, delaying the onset of atherosclerosis,
XX treating or regressing atherosclerosis without inducing
XX hypertriglyceridaemia, in a mammal lacking an endogenous,
XX normally functioning apoE gene or low density lipoprotein (LDL)
XX receptor or is at risk for developing atherosclerosis due to
XX accumulation of lipoprotein remnants in the bloodstream or having
XX a defect in remnant removal.
XX
XX Sequence 1156 BP; 207 A; 369 C; 432 G; 148 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 4,84e-102 Length: 1156
Score: 1370.00 Matches: 276
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 0
Query Match: 99.71% Indels: 0
DB: 24 Gaps: 0
US-09-827-854-15_copy_1_277 (1-277) x AAD22052 (1-1156)
Qy 1 MetLysValLeuTrrPAlaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGTTCTGTGGCTGTGCTGTGTCATTCCTCGCAGAGATCCAGCCAAAGTGT 120
Qy 21 GluGlnAlaValaIuThrGluProGluProGluLeuArgGlnIleThrGluTrrPglInsr 40
Db 121 GAGCAAGCGGTGGAGACAGAGCGGAGCGCCGAGACTCGCGCAGACACCGATGGCAGAGC 180
Qy 41 GlyIleArgTrrPgluLeuAlaLeuGlyArgPheTrrPAspTrrPLeuArgTrrPValGlnThr 60
Db 181 GGGCAGCGCTGGGAACTGCGACCTGGCTTTGGGATTACTGGCGTGGTGGAGACA 240
Qy 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValIleThrGlnLeuArgAla 80
Db 241 CTGCTGAGCAGCTGCAGAGAGAGTGTCTCAGCTCCAGCTCACCCAGAACTGAGGGCG 300
Qy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
Db 301 CTGATGACAGACACCATGAAGGAGTTGAAGCCTACAAATCGGAGACTGGAGCAACTG 360
Qy 101 ThpProValAlaIugIurThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGACGGCGGCGGCTGTCCAAAGAGAGTGCAGAGCGCGGACAGGCC 420
Qy 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
Db 421 CGGCTGGCGCGGACAGATGAGAGAGTGTGGCGCGCTGTGTGCAATCCCGCGGAGAGTG 480
Qy 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGCTCGGCCACAGACACCGAGAGACTCGGGGTGGCGCTCCGCTCCACCTGCGC 540
Qy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCGTAAGCGGCTCTCCCGCATGCCGATGACCTGCAGAAAGCGCCTGCGAGTGTAC 600
Qy 181 GlnAlaGlyAlaArgGlyGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
```

```

|||||
Db 601 CAGCGCGGGCCCGCAGGGCCGCGAGCGGCGCTCAGCGCCATCCGCGCGCCCTGGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGTAACAGAGCGCGCGCTGCGCGCCACACTGTGTGGCTCCCTGGCCGCGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleThrPglIuArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGAGCCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGACCCGCGACCGCTGGAGCGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
Db 841 CTGGAGAGACAGCGCCAGAGATACGCTTCAGAGCGCGGCTTCAGAGCC 891

RESULT 7
AAF84314
ID AAF84314 standard; cDNA; 1156 BP.
XX
AC AAF84314;
XX
XX 21-JUN-2001 (first entry)
XX
DE Human ApoE2 coding sequence.
XX
XX Human; ApoE2; Alzheimer's disease; arteriosclerosis; ss.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 61..1014
FT FT /*tag= a
FT FT /product= "Human ApoE2"
XX
PN JP2001017028-A.
XX
PD 23-JAN-2001.
XX
XX 28-APR-2000; 2000JP-0128919.
XX
XX 06-MAY-1999; 99JP-0125647.
XX
XX PA (MITU ) MITSUBISHI CHEM CORP.
XX
XX DR WPI: 2001-285406/30.
XX
XX DR P-PSDB; AAB80996.
XX
XX PT New apoe humanized mammalian cell useful for screening for agents
XX PT useful for treating or preventing Alzheimer's disease and
XX PT arteriosclerosis -
XX
XX PS Disclosure: Page 11-12; 22pp; Japanese.
XX
XX CC The present invention relates to an ApoE humanised mammalian cell. The
XX CC present sequence is the coding sequence for human ApoE2, which was used
XX CC in the method of the present invention. The ApoE humanised mammalian cell
XX CC can be used for screening for agents useful for treating or preventing
XX CC Alzheimer's disease and arteriosclerosis.
XX
XX SQ Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;

Alignment Scores:
Pred. No.: 1,02e-101
Score: 1366.00
Percent Similarity: 99.64%
Best Local Similarity: 99.64%
Query Match: 99.42%
DB: 22
Gaps: 0

```

```

us-09-827-854-15_copy_1_277 (1-277) x AAF84314 (1-1156)
QY 1 MetLysValLeuThrPheAlaLeuLeuValThrPheLeuAlaGlySerGlnAlaLysVal 20
Db 61 ATGAAAGTTCTGTGGGCTCGCTGTGTGTCACATTCCTGGCAGAGGCCAAGGAGTGG 120
QY 21 GlnGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrPglIns 40
Db 121 GAGCAAGCGGTGAGAGACAGACCGGAGCCCGAGCTCGCCAGCAGACCCAGTGGCAGAG 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTrpLeuArgTrpValGlnThr 60
Db 181 GGCACACCGCTGGAACTGACACTGGCTCGCTTTGGATTAACCTGCGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db 241 CTGTCTGACAGAGTGAGAGAGAGCTGCTCAGCTCCAGATCCAGAGAACTGAGAGGGG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnGlnLeu 100
Db 301 CTGATGACAGACACCATGAAGAGATTGAAGCCCTTACAAATCGGAACCTGAGAGAACTG 360
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTGGCGAGAGAGACGCGGCGACGCTGTCCAAAGAGCTGCAGCGCGCGCAGGCC 420
QY 121 ArgLeuGlyAlaAspMetGluAspValLysGlyArgLeuValGlnTrpArgGlyGluVal 140
Db 421 CGGCTGGGCGGACATGAGAGAGCTGTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 141 GlnAlaMetLeuGlnLysInsThrGlnGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
Db 481 CAGCGCATGCTCGCGCCAGAGACACCGAGAGAGCTCGGGGTGCTGCTGCTGCCACTGGCC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTCGTAAAGCGGCTCTCTCCGATGCCGATGACCTCAGAGAGTGGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlnLysAlaGluArgGlyLeuSerAlaIleArgGlnThrArgLeuGly 200
Db 601 CAGGCGGGGCGCGCGAGGCGCGCGAGCGCGGCTCAGCGCCATCCGCGCAGCGCCGCGGG 660
QY 201 ProLeuValGluGlnGlyArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGTGTGTAACAGAGCGCGCGCTGCGCGCCACACTGTGTGGCTCCCTGGCCGCGCCG 720
QY 221 LeuGlnGluArgAlaGlnAlaIleThrPglIuArgLeuArgAlaArgMetGluGluMetGly 240
Db 721 CTACAGAGAGCGGGCCAGAGCCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
Db 781 AGCCGAGACCCGCGACCGCTGGAGCGAGGTGAAGAGCAGGTGGCGGAGGTGGCGCCAG 840
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla 277
Db 841 CTGGAGAGACAGCGCCAGAGATACGCTTCAGAGCGCGGCTTCAGAGCC 891

RESULT 8
AAD22049
ID AAD22049 standard; DNA; 1156 BP.
XX
AC AAD22049;
XX
XX 12-FEB-2002 (first entry)
XX
XX DE Human apolipoprotein E (apoE) isoprotein, apoE2 DNA.
XX
XX DT Human; apolipoprotein E; apoE; cholesterol; arteriosclerosis;
XX KW hypertriglyceridaemia; low density lipoprotein; LDL; ds.
XX
XX OS Homo sapiens.
XX

```

FH	Key	Location/Qualifiers
FT	CDS	61..1014
FT		/*tag= a
FT		/product= "Human apoE isoprotein, apoE2"
FT	sig_peptide	61..114
FT		/*tag= b
FT	mat_peptide	115..1011
FT		/*tag= c
FT		/product= "Mature human apoE isoprotein, apoE2"
XX		
PN		WO200177136-A1.
XX		
PD		18-OCT-2001.
XX		
PF		06-APR-2001; 2001WO-US11358.
XX		
PR		06-APR-2000; 2000US-0544386.
PR		04-OCT-2000; 2000US-0679088.
PR		05-APR-2001; 2001US-0827854.
XX		
PA		(KOSP-) KOS PHARM INC.
PA		(UYBO-) UNIV BOSTON.
PI		Zannis VI, Kypros KE;
XX		
DR		WPI: 2002-010885/01.
XX		P-PSDB; AAEI3295.
XX		
PT		New apolipoprotein E polypeptide and nucleic acid, useful for lowering
PT		cholesterol, delaying the onset of or treating atherosclerosis in
PT		mammal, without inducing hypertriglyceridemia
XX		
XX		Claim 14; Page 81-82; 91pp; English.
XX		
CC		The present sequence is a human apolipoprotein E (apoE)
CC		isoprotein, apoE2 DNA. The apoE lipoproteins are useful for
CC		lowering cholesterol, delaying the onset of atherosclerosis,
CC		treating or regressing atherosclerosis without inducing
CC		hypertriglyceridemia, in a mammal lacking an endogenous,
CC		normally functioning apoE gene or low density lipoprotein (LDL)
CC		receptor or is at risk for developing atherosclerosis due to
CC		accumulation of lipoprotein remnants in the bloodstream or having
CC		a defect in remnant removal.
XX		
XX		
SQ		Sequence 1156 BP; 208 A; 367 C; 432 G; 149 T; 0 other;
	Alignment Scores:	
	Pred. NO.:	1,02e-101
	Score:	1366.00
	Length:	1156
	Matches:	276
	Conservative:	0
	Mismatches:	1
	Indels:	0
	Gaps:	0
US-09-827-854-15_COPY_1_277 (1-277) x AMD22049 (1-1156)		
OY	1	MethylsValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
Db	61	ATGAAGGTTCTGTGGCTGGCTGGTCTGGTCACTTCCTGGAGATGCCAGGCCAAAGTG 120
OY	21	GlnGlnAlaValaGlnThrGlnProGlnProGlnLeuAlaGlnGlnThrGlnTrpGlnSer 40
Db	121	GAGCAACCGGTGGAGACAGAGCCGGAGCCCGACTCGCCACAGCCAGGCGCAGAGC 180
OY	41	GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
Db	181	GGCCAGCGCTGGGAACCTGGCACTGGGTGGCTTTGGGATTTACCTGGGCTGGAGACA 240
OY	61	LeuSerGlnGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
Db	241	CTGTCTAGACAGGTGACAGAGAGAGAGCTGTACGTCCACAGTACCACCCAGAACTAGAGGCG 300
OY	81	LeuMetAspGluThrMetLysGlnLeuLysAlaTyrLysSerGlnLeuGlnGlnLeu 100

Db	301	CGATGGAGCAGACCATGAGAGATTGAAGGCTCAAAATGTGGAACCTGGAGAACCACTG	360
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerIysGluLeuGlnAlaAlaGlnAla	120
Db	361	ACCCGCGTGGCGGAGAGAACGCGGGCAGCGCTGTCCAAAGCAGCTGACAGCGCGCAGGCC	420
QY	121	ArgLeuGluAlaAspMetGluAspAlaCysGluIlyrArgLeuValGlnIlyrArgIlyVal	140
Db	421	CGGCTGGCGCGCAGATGAGACCTGTGGCGCGCCCTGGTGTGCAGTACCGCGGCGAGGTG	480
QY	141	GlnAlaMetLeuGluGlnSerThrGluIleuIlyrArgValArgLeuAlaSerHisLeuArg	160
Db	481	CAGGCGATGCTCTGGCCAGACACCGAGAGCTGGCGGGTCTGGCTCTCCACTCTGCC	540
QY	161	LysLeuArgIlyAspArgLeuLeuArgAspAlaAspAspLeuGlnIlyrArgLeuAlaValTyr	180
Db	541	AAGCTGCGTAAAGCGGCTCTCTCCGCGATGCCGATGACCTGCAGAGAGTGCCTGGCAGTGTAC	600
QY	181	GlnAlaGluAlaArgGluGluAlaGluIlyrArgIlyLeuSerAlaIleArgGluIlyrLeuGly	200
Db	601	CAGCGCGGGGGCGCGGAGGGCGCGCAGCGCGCTGCAGAGCGCAATCCGCGAGCGCTGGGG	660
QY	201	ProLeuValGluGlnGlyArgValArgAlaIleAlaThrValGlySerLeuAlaGlyGlnPro	220
Db	661	CCCCCTGGTGGAAACAGGGCGCGCTGGCGGCCGCACCTGTGGGCTCTCCCTGGCGGCGACCG	720
QY	221	IeuGlnGluIlyrArgAlaGlnAlaIleTyrGlyIlyrArgLeuArgAlaArgMetGluGluMetGly	240
Db	721	CTACAGAGAGCGGGGCCAGGCGCTGGGGGCGAGCGGCTGGCGCGCGCGGATGAGAGAGATGGGC	780
QY	241	SerArgThrArgAspArgIlyLeuAspGluValIlysGluGlnValAlaGluValArgAlaIlys	260
Db	781	AGCCGAGACCCGCGACCGCTGTGACGAGGTGAAGAGCAGAGTGGGAGGTGGCGGCCAAG	840
QY	261	IeuGlnGluGlnAlaGlnGlnIleArgArgGlnAlaGluAlaPheGlnAla	277
Db	841	CTGGAGGAGCAGAGCCACGACATACGCTGTGAGGCCGAGGCTTCCAGGCC	891
RESULT 9			
AAID22051	ID	AAID22051 standard; DNA; 1156 BP.	
AC	AAID22051;		
XX	12-FEB-2002	(first entry)	
DE	Human apolipoprotein E (apoE) allele, apoE2* DNA.		
KM	Human; apolipoprotein E; apoE; cholesterol; atherosclerosis;		
XX	hypertriglyceridaemia; low density lipoprotein; LDL; ds.		
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	61..1014	
FT		/*tag= a	
FT	sig_peptide	/product= "Human apoE allele, apoE2*"	
FT		61..114	
FT		/*tag= b	
FT	mat_peptide	115..1011	
FT		/*tag= c	
FT		/product= "Mature human apoE allele, apoE2*"	
XX	WO20017136-A1.		
PD	18-OCT-2001.		
XX			
PE	06-APR-2001; 2001WO-US11358.		
PR	06-APR-2000; 2000US-0544386.		
PR	04-OCT-2000; 2000US-0679088.		
PR	05-APR-2001; 2001US-0827854.		

CC used in therapeutic applications.

XX Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	1,23e-101	Length:	1157
Score:	1365.00	Matches:	275
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.34%	Indels:	0
	17	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x AAT06957 (1-1157)

```

QY 1 MetlyValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
   |||||
Db 62 ATGAGGTTCTGTGGCTGGTGTGCTGTCACATCTCTGGCAGATGCCAGGCCAAGT 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
   |||||
Db 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCGCGCAGACAGACCGAGTGCAGAGC 181
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
Db 182 GGCAGCGCTGGGAACTGGCACTGGGTGCTTTGGATTACCTGGCTGGGTGCAGACA 241
QY 61 LeuSerGluGlnValGlnGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
   |||||
Db 242 CTGCTGAGCAGAGTGAGGAGAGAGCTGCTCAGCTCCCAATCCACCAAGAACTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
   |||||
Db 302 CTGATGACAGACCAATGAGGAGATTGAAAGCCTACAAATCGGAAGTGGAGAACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
   |||||
Db 362 ACCCGGTTAGCGAGAGAGACCGGGCCAGCTGTCCAAAGACTGCGAGAGCGCGCAGGCC 421
QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnThrArgGlyGluVal 140
   |||||
Db 422 CGGCTGGCGCCGACATGAGAGAGAGCTGTGGCGCGCTGTGCAGTACCCGCGGAGAGTG 481
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
   |||||
Db 482 CAGGCGCATGCTCGCCACAGACACCGAGAGCTCGGGTGGCGCTCCGCTCCACCTGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
   |||||
Db 542 AAGCTGGGTAAAGCGGCTCTCCCGGATCCCGATGACTCTGAGAAAGCCCTGGCAGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
   |||||
Db 602 CAGCGCGGGGCGCGGAGGCGCGGAGCGCGGCTCAGCCATCCGCGAGCGGCTGGGG 661
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValLysSerLeuAlaGlyGlnPro 220
   |||||
Db 662 CCCCTGGTGAAGAGGCGCGCGTGGCGCGCGCAGCTGTGGCTCCCTGGCGCGCGCACCGC 721
QY 221 LeuGlnGluArgAlaGlnAlaTArgLysArgLeuArgAlaArgMetGluGluMetGly 240
   |||||
Db 722 CTACAGAGAGCGGCGCCAGGCGCTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
   |||||
Db 782 AGTCGAGCCCGGAGCGCGCTGGAGCAGAGTGAAGAGAGAGTGGCGGAGTGGCGCGCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
   |||||
Db 842 CTGAGGAGAGCAGGCCAGCAGATACGCTTGCAGGCGAGGCCCTTCCAGAGGCC 892

```

RESULT 11
ABN95746
ID ABN95746 standard; DNA: 1157 BP.

```

XX AC ABN95746;
XX XX 13-AUG-2002 (first entry)
DT DE Gene #2244 used to diagnose liver cancer.
XX XX
XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX KM metastatic liver tumour; cytostatic; expression profile; disease state;
XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX OS Homo sapiens.
XX PN WO200229103-A2.
XX PD 11-APR-2002.
XX PF 02-OCT-2001; 2001WO-US30589.
XX PR 02-OCT-2000; 2000US-237054P.
XX PA (GENE-) GENE LOGIC INC.
PI Horne D, Alvares C, Petes-Da-Silva S, Vockley JG;
XX MPI: 2002-426119/45.
XX DR
XX PT diagnosing and detecting the progression of liver cancer,
XX PT hepatocellular carcinoma or metastatic liver tumor in a patient,
XX PT involves detecting the level of expression of two or more genes in a
XX PT liver tissue sample
XX PS Claim 1; SEQ ID NO 2244; 298bp; English.
XX XX
XX CC The invention relates to a novel method for diagnosing and detecting the
XX CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX CC tumour in a patient, and differentiating metastatic liver cancer from
XX CC hepatocellular carcinoma in a patient, involving detecting the level of
XX CC expression of two or more genes represented in ABN93503-ABN97455 in a
XX CC tissue sample. The method of the invention has hepatotropic, and
XX CC cytostatic activity. The method is useful for diagnosing and detecting
XX CC the progression of liver cancer, hepatocellular carcinoma and metastatic
XX CC liver carcinoma in a patient. The method is useful for identifying
XX CC expression profiles which serve as useful diagnostic markers as well as
XX CC markers that can be used to monitor disease states, disease progression,
XX CC drug toxicity, drug efficacy and drug metabolism.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other;

```

Alignment Scores:

Pred. No.:	1,23e-101	Length:	1157
Score:	1365.00	Matches:	275
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.34%	Indels:	0
	24	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x ABN95746 (1-1157)

```

QY 1 MetlyValLeuTrpAlaLeuValThrPheLeuAlaGlyCysGlnAlaVal 20
   |||||
Db 62 ATGAGGTTCTGTGGCTGGTGTGCTGTCACATCTCTGGCAGATGCCAGGCCAAGT 121
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
   |||||
Db 122 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTCGCGCAGACAGACCGAGTGCAGAGC 181
QY 41 GlyAlaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
   |||||
Db 182 GGCAGCGCTGGGAACTGGCACTGGGTGCTTTGGATTACCTGGCTGGGTGCAGACA 241

```

QY 61 LeuSerGluGlnValGlnGlnGlnLeuLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||
 Db 242 CTGTCTGACAGAGTGCAGAGAGAGTGTCTCAGTCCAGTCCCAAGTCCAGCAAGACTGAGGGCG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 |||
 Db 302 CTGATGAGACAGACCATGTAAGGAGTGTGAAGGCTTACAAATCGGAATCGAGAGAACAACTG 361
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 |||
 Db 362 ACCCGGTAGCGAGAGAGAGCGGGGACGCTGTCCAAAGACTGAGAGCGCGCAGCGCC 421
 QY 121 ArgLeuGlnAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 |||
 Db 422 CGGCTGGGGCGGACATGAGAGACGTGTCCGCGCTGTGTGACAGTCCGCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlnLysInSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 |||
 Db 482 CAGGCCATGCTCGGCCAGACACCGAGAGCTGCGGGTCCGCTCGCTCCACCTGCGCG 541
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||
 Db 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCCGATGACTGACAAAGCGCTGCGCAGTGTAC 601
 QY 181 GlnAlaGlyAlaArgGlnGlnValArgGlyLeuSerAlaIleArgGlnArgLeuGly 200
 |||
 Db 602 CAGCGCGGGCGCGGAGGCGCGGCGGCGGCTCAGCGCCATCCGCAAGCGCTGAGG 661
 QY 201 ProLeuValGlnGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
 |||
 Db 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGGCGGCTGTGGCTCCCTGCGCGGACCGG 721
 QY 221 LeuGlnGlnLysArgAlaGlnAlaThrPheGlyGlnArgLeuArgAlaArgMetGlnGlnMetGly 240
 |||
 Db 722 CTACAGAGAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 781
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaLys 260
 |||
 Db 782 AGTCGAGACCGCGGACCGCTGTGACGAGGTGAAGGAGACAGAGTGGCGGCGGCGGCGG 841
 QY 261 LeuGlnGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 |||
 Db 842 CTGAGAGAGACGAGCCGAGCATACGCTGAGCGCGGAGCGCTTCCAGGCC 892
 RESULT 12
 ABRK64514
 ID ABRK64514 standard; DNA; 1157 BP.
 AC ABRK64514;
 XX
 XX 18-JUN-2002 (first entry)
 DT
 XX
 DE Human benign prostatic hyperplasia gene #409.
 KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
 OS Homo sapiens.
 XX
 XX WO200212440-A2.
 PN
 XX
 PD 14-FEB-2002.
 XX
 PF 07-AUG-2001; 2001WO-US24708.
 XX
 PR 07-AUG-2000; 2000US-223323P.
 PR 05-JUN-2001; 2001US-0873319.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA (NISB) JAPAN TOBACCO INC.
 XX
 PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 XX

DR WPI: 2002-257476/30.
 XX
 PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells
 PS
 XX
 XX Disclosure: Page 239-240; 44pp; English.
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABRK64106-ABRK64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.
 XX
 SQ Sequence 1157 BP; 212 A; 370 C; 426 G; 149 T; 0 other:
 Alignment Scores:
 Pred. No.: 1,23e-101 Length: 1157
 Score: 1365.00 Matches: 275
 Percent Similarity: 99.28% Conservative: 0
 Best Local Similarity: 99.28% Mismatches: 2
 Query Match: 99.34% Indels: 0
 DB: 24 Gaps: 0
 US-09-827-854-15_COPY_1_277 (1-277) x ABRK64514 (1-1157)
 QY 1 MethylsValLeuThrPheAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||
 Db 62 ATGAAGGTTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 121
 QY 21 GlnGlnAlaValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
 |||
 Db 122 GAGCAAGCGGTGAGACAGAGCGGAGCCGAGCTGCGCCAGACAGACCGAGTGGCAGAGC 181
 QY 41 GlnGlnArgTyrPheGluLeuAlaLeuGlnLysArgPheTyrPheTyrLeuArgTyrValGlnThr 60
 |||
 Db 182 GGGCAGCGCTGGAGACTGGACACTGGGTGCTTTGGATTAACCTGCGCTGGGTGAGACACA 241
 QY 61 LeuSerGlnGlnValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
 |||
 Db 242 CTGTCTGAGACAGGTGAGAGAGGAGTGTCACTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 301
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGlnLeuGlnGlnAlaGlnAla 100
 |||
 Db 302 CTGATGAGACAGACCATGTAAGGAGTGTGAAGGCTTACAAATCGGAATCGAGAGAACAACTG 361
 QY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaAlaGlnAla 120
 |||
 Db 362 ACCCGGTAGCGAGAGAGAGCGGGGACGCTGTCCAAAGACTGAGAGCGCGCAGCGCC 421
 QY 121 ArgLeuGlnAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
 |||
 Db 422 CGGCTGGGGCGGACATGAGAGACGTGTCCGCGCTGTGTGACAGTCCGCGCGGAGGTG 481
 QY 141 GlnAlaMetLeuGlnLysInSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160


```
Db 182 GGCACACCGCTGGGAACCTGCACCTGGCTCTTTTGGATTACCTGCCTGGGTGCAGACA 241
QY LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
Db 242 CTGTCTGACGACGATGAGAGAGCTGCTCAGCTCCCAAGTACCCAAAGAACTGAGGGCG 301
QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeu 100
Db 302 CAGATGAGACGACCATGAGAGAGTTGAAGGCTTCAAAATCGGAATCGAGAGACACACG 361
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
Db 362 ACCCGGTGACGAGAGACGCGGCGACGCGTCTCAAGAGACTGCAGACGCGCAGGCG 421
QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGIuTyrArgGIuVal 140
Db 422 CGGCTGGGGCGGACATGAGAGAGCTGTGCGGCCCTGTGCTACATCCGCGGCAAGGTG 481
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
Db 482 CAGGCGCATGCTCGGCGACAGACCGAGAGCTGCGGGTGGCTGCGCTGCCCACTGGCGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
Db 542 AAGCTGCGTAAACGGGCTCTCTCCGATGCCGATGACCTGCAGAAAGCGCTGGCAGGTGAC 601
QY 181 GlnAlaGIuAlaArgGIuGIuValArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
Db 602 CAGGCGCGGGGGCCCGGAGGGCGCCGAGCGCGGCTCAGCGCATCCGCGAGCGGCTGGGG 661
QY 201 ProLeuValGIuGIuGIuArgValArgValArgAlaThrValGIuSerLeuAlaGIuInPro 220
Db 662 CCCCTGGTGGAAACAGAGGGCGGCTGGGGCGGCCACTGTGGGTCTCCCTGGCGGCCAGCGG 721
QY 221 LeuGIuGIuArgAlaGIuAlaThrPcLysGIuArgLeuArgAlaArgMetGIuGIuMetGIu 240
Db 722 CTACAGAGAGCGGCGCCAGGCGCTGGGGCGAGCGGCTGCCGCGGATGGAGAGATGGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGIuValLysGIuGIuValAlaGIuValArgAlaLys 260
Db 782 ACTCGAGACCGCGACCGCTGGACGAGGAGAGAGAGGTGGGAGGTGGCGCCCAAG 841
QY 261 LeuGIuGIuGIuAlaGIuGIuIleArgLeuGIuAlaGIuAlaPheGIuAla 277
Db 842 CTGAGAGAGACGAGCCGACAGATAGCTGCGAGGCGGAGGCGCTTCAGAGCC 892

RESULT 14
AAN50450
ID AAN50450 standard; DNA; 1110 BP.
XX
AC AAN50450;
XX
DT 09-JAN-1992 (first entry)
XX
DE Sequence encoding human apolipoprotein E.
XX
KW Hyperlipaemia; ds.
XX
OS Homo sapiens.
XX
FH Key 15..968 Location/Qualifiers
FT CDS /*tag= a
FT mat_peptide 69..965
FT /*tag= b
XX
XX JP60118189-A.
XX
XX 25-JUN-1985.
XX
XX 29-NOV-1983; 83JP-0224980.
XX
XX 29-NOV-1983; 83JP-0224980.
```

```
XX
PA (MITU ) MITSUBISHI CHEM IND KK.
XX
DR WPI: 1985-188003/31.
DR P-PSDB: AAP51204.
XX
PT DNA fragment - contg. DNA which codes human apolipoprotein E for
PT treatment of hyperlipaemia.
XX
PS Claim 3; Page 484; 8pp; Japanese.
XX
CC The sequence may be used to produce the apolipoprotein E, useful in the
CC treatment of hyperlipaemia.
XX
SQ Sequence 1110 BP; 198 A; 355 C; 414 G; 143 T; 0 other.

Alignment Scores:
Pred. No.: 1,41e-101 Length: 1110
Score: 1364.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.27% Indels: 0
DB: 6 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AAN50450 (1-1110)
QY 1 MetLysValIleuThrPAlaIleuLeuValThrPheLeuAlaGIuGIuAlaLysVal 20
Db 15 ATGAAGGTTCTGTGGGCTGGCTGTGCTGACATTCCTGCGAGATGCCAGGCAAGGTG 74
QY 21 GlnAlaIleuAlaGIuThrGIuProGIuProGIuLeuArgGIuGIuThrGIuTrpGIuSer 40
Db 75 GAGCAAGCGGTGGAACAGACAGCGGAGCCCGAGCTGCGCCACAGACCGAGCGAGACG 134
QY 41 GlyIuArgTrpGIuLeuAlaLeuGIuLysArgPheTrpAspTyrLeuArgTrpValGIuThr 60
Db 135 GGCACAGCGCTGGGAACCTGACATGGTGGCTTTTGGATTAACCTGCGCTGGGTGCAGACA 194
QY 61 LeuSerGIuGIuValGIuGIuLeuLeuSerSerGIuValThrGIuGIuLeuArgAla 80
Db 195 CTGTGTGAGCACGTGCAGAGAGAGCTGCTCAGCTCCAGGTACCCAGAGAACTGAGGGCG 254
QY 81 LeuMetAspGIuThrMetLysGIuLeuLysAlaTyrLysSerGIuLeuGIuGIuLeuIn 100
Db 255 CTGATGAGACGAGACCATGATGAGAGAGTTGAAGGCTTCAAAATCGGAATCGAGAGAACTG 314
QY 101 ThrProValAlaGIuGIuThrArgAlaArgLeuSerLysGIuLeuGIuAlaGIuAla 120
Db 315 ACCCGGTGGCGAGAGAGCGGGCGACGGCTGTCTCAAGAGACTGCAGGCGCGCCAGGCGC 374
QY 121 ArgLeuGIuAlaAspMetGIuAspValLysGIuArgLeuValGIuTyrArgGIuVal 140
Db 375 CGGCTGGCGCGGACATGAGAGAGCTGTGCGGCCCTGTGGTCACTATACCAGCGGCAAGGTG 434
QY 141 GlnAlaMetLeuGIuGIuSerThrGIuGIuLeuArgValArgLeuAlaSerHisLeuArg 160
Db 435 CAGGCGCATGCTGGCGACAGACCGAGAGAGCTGCGGGTGGCTGCCCTCCCACTGGCGC 494
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGIuLysArgLeuAlaValTyr 180
Db 495 AAGCTGCTAAAGCGGCTCTCTCCGATGCCGATGACCTGCAGAAAGCGCTGCACTGTAC 554
QY 181 GlnAlaGIuAlaArgGIuGIuValArgGIuArgGIuLeuSerAlaIleArgGIuArgLeuGIu 200
Db 555 CAGGCGGGGGCGCGGAGAGCGCGAGGCGGCTCAGCGCCATCGGAGAGGCTTGGGG 614
QY 201 ProLeuValGIuGIuGIuArgValArgValArgAlaThrValGIuSerLeuAlaGIuInPro 220
Db 615 CCCCTGGTGGAAACAGAGGCGGCTGGGGCGGCACTGTGGGTCTCCCTGGCGGCCAGGCGG 674
QY 221 LeuGIuGIuArgAlaGIuAlaThrPcLysGIuArgLeuArgAlaArgMetGIuGIuMetGIu 240
Db 675 CTACAGAGAGCGGCGCCAGGCGCTGGGGCGAGGCGGCTGCCGCGGATGAGAGATGGGC 734
```

QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 735 AGCGGACCGCGACCGCCGTGAGAGAGTGAAGAGAGGTGGCGGAGGTGGCGCCAG 794
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 795 CTGGAGAGAGCGCCAGCATACGCTGACGCGCCAGGCGCTTCCAGGCC 845

RESULT 15
AAF84316
ID AAF84316 standard; cDNA; 1156 BP.
AC AAF84316;
DT 21-JUN-2001 (first entry)
DE Human ApoE4 coding sequence.
KW Human; ApoE4; Alzheimer's disease; arteriosclerosis; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 61..1014
FT /tag= a
FT /product= "Human ApoE4"
XX
XX JP2001017028-A.
XX
XX 23-JAN-2001.
XX
XX 28-APR-2000; 2000JP-0128919.
XX
XX PR 06-MAY-1999; 99JP-0125647.
XX
XX (MITU) MITSUBISHI CHEM CORP.
XX
XX PA WPI: 2001-285406/30.
XX DR P-PSDB; AAB80998.
XX
XX New apoe humanized mammalian cell useful for screening for agents
PT useful for treating or preventing Alzheimer's disease and
PT arteriosclerosis -
XX
XX PS Disclosure; Page 16-17; 22pp; Japanese.
XX
XX CC The present invention relates to an ApoE humanised mammalian cell. The
CC present sequence is the coding sequence for human ApoE4, which was used
CC in the method of the present invention. The ApoE humanised mammalian cell
CC can be used for screening for agents useful for treating or preventing
CC Alzheimer's disease and arteriosclerosis.
XX
XX SO Sequence 1156 BP; 208 A; 369 C; 432 G; 147 T; 0 other;

Alignment Scores:
Pred. No.: 2.15e-101 Length: 1156
Score: 1362.00 Matches: 276
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.13% Indels: 0
DB: 22 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x AAF84316 (1-1156)

QY 1 MetLysValIleuTrpAlaIleuLeuValThrPheLeuAlaGlySerGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGGCTGTGTGTCACATTCCTGCGAGATGCCAGGCCAAGGTG 120
QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 121 GAGCAAGCGGTGAGACAGAGCCGAGCCCGAGCTGCCACAGACCCGAGTGGCAGAGC 180

QY 41 GlyGlnArgTrpGluLeuAlaIleuGlnGlyArgPheTrpAspTyrIleuArgTrpValGlnThr 60
DB 181 GGCCAGCGGCTGGGAACTGGCACTGGGTTCGCTTTGGGATTACCTGGCTGGGTGCAGACA 240
QY 61 LeuSerGluGlnValGlnGluGlnIleuLeuSerSerGlnValThrGlnGlnIleuArgAla 80
DB 241 CTGTCTGAGCAGGTGCAGAGAGAGCTGCTCAGCTCCAGAGTCAACCAGAACTGAGAGGCC 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnGlnIleu 100
DB 301 CTGATGGACAGACCATGAAGAGTTGAAGGCTTACAAATCGGAACCTGAGAGAACTAG 360
QY 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTGGCGAGGAGAGACGCGGCGACGCTGTCCAGAGAGCTGCACGAGCGCGCAGGCC 420
QY 121 ArgLeuGlyValAspMetGluAspValLysGlyArgLeuValGlnThrArgGlyGluVal 140
DB 421 CGGCTGGCGCGGACATGGAGAGCTGCGCGGCGCTGTGTCAGTACCGGCGAGGTG 480
QY 141 GlnAlaMetLeuGlnGlnSerThrGluGluIleuArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGGCCAGAGCACCGAGAGCTCGGGTGGCTCGCTGCCCTGCCACTGCGCG 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AACCTGCTAAGCGGCTCCTCCGCGATGCCGATGACCTGAGAAAGCGCTGGCAGGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGCGCGGGCGCGGAGGCGCGGAGCGGCGCTCAGCGCCATCCGCGAGCGCTGGGG 660
QY 201 ProLeuValGluGlnGlnLysArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 661 CCCCTGTGGAAACAGGCGCGCGCTGGCGGCGCGCACCTGTGGGCTCCTGGCGCGCGCAGCGG 720
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyGluArgLeuArgAlaArgMetGluMetGly 240
DB 721 CTACAGAGAGCGGGCCAGGCTGTGGGCGAGCGGCTGCGCGCGGATGAGAGATGGGC 780
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 781 AGCGGAGCCCGCACCGCTGGACGAGGTGAAGAGCAGGTGGAGGTGGCGGCCAAG 840
QY 261 LeuGluGlnGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 841 CTGGAGAGAGCGCCAGCATACGCTGACGCGCCAGGCGCTTCCAGGCC 891

Search completed: March 14, 2003, 12:43:02
Job time : 161.176 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 12:17:52 ; Search time 30.3514 Seconds
(without alignments)
2798.866 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Perfect score: 1374
Sequence: 1 MKVLMALVTLFLAGCCAKY.....RAKLEGAQOIRLOAEAFQA 277

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09827554/rumat_11032003_101611_27510/app-query.fasta.1.3576
-DB=issued.patents.NA -OPT=fastap -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09827554 @CGN.1.1.224 @rumat_11032003_101611_27510 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGESOURCY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=130
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
Issued.patents.NA.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/6D.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6E.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1365	99.3	1157	1	US-07-709-949-1 Sequence 1, Appli
2	1002.5	73.0	1126	4	US-08-949-155-5 Sequence 5, Appli
3	1002.5	73.0	1126	4	US-09-819-964-5 Sequence 5, Appli
4	984	71.6	660	2	US-08-726-106A-28 Sequence 28, Appli
5	866.5	63.1	4267	4	US-08-949-155-51 Sequence 51, Appli
6	866.5	63.1	4267	4	US-09-819-964-51 Sequence 51, Appli
7	511	37.2	330	1	US-07-849-389-6 Sequence 6, Appli
8	374	27.2	252	3	US-08-617-256-24 Sequence 24, Appli
9	374	27.2	252	4	US-09-287-141-24 Sequence 24, Appli
10	374	27.2	252	4	US-09-431-613-24 Sequence 24, Appli
11	374	27.2	252	4	US-09-504-245-24 Sequence 24, Appli
12	374	27.2	252	4	US-09-287-682-24 Sequence 24, Appli

13	374	27.2	252	4	US-09-287-679-24 Sequence 24, Appli
14	374	27.2	252	4	US-09-397-766-24 Sequence 24, Appli
15	374	27.2	252	4	US-09-287-681-24 Sequence 24, Appli
16	374	27.2	252	4	US-09-495-444-24 Sequence 24, Appli
17	161.5	11.8	801	1	US-07-939-946-4 Sequence 4, Appli
18	161.5	11.8	801	1	US-08-333-577-4 Sequence 4, Appli
19	161.5	11.8	801	5	PCT-US92-08634-4 Sequence 4, Appli
20	161.5	11.8	842	4	US-08-938-105-2 Sequence 2, Appli
21	152.5	11.1	5661	4	US-08-938-105-2 Sequence 2, Appli
22	147	10.7	1879	4	US-09-750-580-2 Sequence 2, Appli
23	146.5	10.7	3256	2	US-08-968-751-3 Sequence 3, Appli
24	141.5	10.3	863	1	US-08-448-606-7 Sequence 7, Appli
25	141.5	10.3	964	1	US-08-448-606-5 Sequence 5, Appli
26	141	10.3	6306	1	US-08-195-487-3 Sequence 3, Appli
27	141	10.3	6306	5	PCT-US93-06160-3 Sequence 3, Appli
28	140	10.2	11236	1	US-07-853-913-1 Sequence 1, Appli
29	137.5	10.0	516	4	US-09-183-861-34 Sequence 34, Appli
30	137.5	10.0	516	4	US-09-022-765-34 Sequence 34, Appli
31	137.5	10.0	13121	4	US-08-961-527-126 Sequence 126, App
32	137	10.0	6306	1	US-08-466-390-3 Sequence 3, Appli
33	137	10.0	6306	1	US-08-470-950-3 Sequence 3, Appli
34	137	10.0	6306	1	US-08-467-781-3 Sequence 3, Appli
35	137	10.0	6306	2	US-08-483-924-3 Sequence 3, Appli
36	136	9.9	1771	2	US-08-533-669A-7 Sequence 7, Appli
37	136	9.9	1771	2	US-08-511-872-1 Sequence 1, Appli
38	136	9.9	1771	4	US-09-183-861-7 Sequence 7, Appli
39	136	9.9	1771	4	US-09-022-765-7 Sequence 7, Appli
40	135.5	9.9	4296	4	US-09-060-410-3 Sequence 3, Appli
41	134.5	9.8	603	4	US-08-957-796-14 Sequence 14, Appli
42	132.5	9.6	2085	2	US-08-668-128B-7 Sequence 7, Appli
43	132.5	9.6	2085	2	US-08-905-445-7 Sequence 5, Appli
44	131.5	9.6	8789	1	US-08-328-254-5 Sequence 5, Appli
45	131.5	9.6	10136	1	US-08-353-700-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-07-709-949-1
Sequence 1, Application US/07709949
Patent No. 5,472,858
GENERAL INFORMATION:
APPLICANT: Attie, Alan D
APPLICANT: Gretch, Daniel G
APPLICANT: Sturley, Stephen L
APPLICANT: Beckage, Nancy E
TITLE OF INVENTION: Production of Recombinant Proteins in
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Quarles & Brady
STREET: P.O. Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/709,949
FILING DATE: 19910604
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 9629691801
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 1:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1157 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1015
FEATURE:
NAME/KEY: misc.signal
LOCATION: 62..119
PUBLICATION INFORMATION:
AUTHORS: McLean, J W
JOURNAL: Journal of Biological Chemistry
VOLUME: 259
PAGES: 6498-6504
DATE: 1984
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 1157
US-07-709-949-1

Alignment Scores:
Pred. No.: 9,616-125 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.288 Conservative: 0
Best Local Similarity: 99.288 Mismatches: 2
Query Match: 99.348 Indels: 0
DB: 1 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-07-709-949-1 (1-1157)
QY 1 MetLysValLeuTrpAlaIaIaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAGGTTCTGTGGCTGCTGTTGCTGTACATTCTGCGAGATGCGAGCCAGAGTG 121
QY 21 GUGlnAlaValAGluThrProGluProGluLeuArgGlnGluThrGluTrpGlnSer 40
DB 122 GAGCAAGCGGTGAGACAGACGCGGAGCCGAGCTGCGCAGACAGCCAGTGCGAGGC 181
QY 41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GGCAGCGCTGGGAAGTGGCTGCTGCTTTGGATTACTGGCGCGGCGGCGAGACA 241
QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTCTGACACAGGTGCGAGGAGAGCTGCTCAGCTCCAGATCACCAGAACGAGGCGG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
DB 302 CTGATGACAGACACCATGAGAGGTGAAGGCTTACAAATCGAATCGAGAGAACACTG 361
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 362 ACCCGGATGCGGAGAGAGCGGCGGCGAGCGCTGTCCAAAGAGACTGAGACGCGCGAGCC 421
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
DB 422 CGGCTGGCGCGGACATGAGAGACGTGTGCGCGCTGTCAGTACCGCGGCGAGGTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGCGCAGAGACCGAGGAGCTGCGGGTGGCTGCCCTCCACCTGCGGC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGTGCGCTAACAGCGGCTCTCCGCGATCCCGATGACCTCAGAGAGCGCTGGCAGGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGlnArgLeuGly 200

```

```

DB 602 CAGCGCGGCGCCCGCAGAGGCGCCGAGCGCGCTCAGCGCCATCCGAGGCGCTGGGG 661
QY 201 ProLeuValGluGlnGlyValArgValArgAlaIaIaThrValGlySerLeuAlaGlyLnaPro 220
DB 662 CCCCTGTGTGAACAGGCGCGCTGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 221 LeuGlnGluArgAlaGlnAlaIaTrpGlyGluArgLeuArgAlaArgMetGluLnaMetGly 240
DB 722 CTACAGAGAGCGGCGCCAGGCTGCGGCGAGCGGCTGCGCGCGCGATGAGAGATGGGC 781
QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaLys 260
DB 782 AGTCGACCGCGGACCGCTGAGCAGAGTGAGAGAGCAGTGGCGAGGCTGCGCGCCCAAG 841
QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 842 CTGAGAGAGCAGGCGCCAGCATACGCTGACGCGCGAGCGCTTCCAGGCC 892

RESULT 2
US-08-949-155-5
Sequence 5, Application US/08949155
Patent No. 6271436
GENERAL INFORMATION:
APPLICANT: Piedrahita, Jorge A
APPLICANT: Bazer, Fuller W
TITLE OF INVENTION: Compositions and Methods for the
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESS: ARNOLD, WHITE AND DURKEE
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,155
FILING DATE: Concurrently Herewith
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/027,338
FILING DATE: 11-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/046,094
FILING DATE: 09-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAWK-177
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1126 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 51..1001
US-08-949-155-5

Alignment Scores:
Pred. No.: 2,596-89 Length: 1126
Score: 1002.50 Matches: 198
Percent Similarity: 83.93% Conservative: 37

```

```

1 ADDRESS: ARNOLD, WHITE AND DURKEE
2 STREET: P.O. Box 4433
3 City: Houston
4 STATE: TX
5 COUNTRY: US
6 ZIP: 77210-4433
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: floppy disk
10 COMPUTER: IBM PC compatible
11 OPERATING SYSTEM: PC-DOS/MS-DOS
12 SOFTWARE: PatentIn Release #1.0, Version #1.30
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/09/819,964
16 FILING DATE: 28-Mar-2001
17 CLASSIFICATION: <Unknown>
18
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: 08/949,155
21 FILING DATE: <Unknown>
22 APPLICATION NUMBER: US 60/046,094
23 FILING DATE: 09-MAY-1997
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Hibler, David W.
26 REGISTRATION NUMBER: 41,071
27 REFERENCE/DOCKET NUMBER: TMMK.177
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (512) 418-3000
30 TELEFAX: (713) 789-2679
31
32 INFORMATION FOR SEQ. ID NO: 5:
33
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 1126 base pairs
36 TYPE: nucleic acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39
40 FEATURE:
41
42 NAME/KEY: CDS
43 LOCATION: 51..1001
44
45 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
46
47 US-09-819-964-5
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
```

OY	118	AlaGlnAlaIaArgLeuGluYAlaIaAspMetGluAspValLysGluIaArgLeuValGlnTyrArg	137
Db	399	GGCGAGGCGCCGCGTGGGCGCCGACATGAGAGACTGGCCACCGCTTGGTGTCTTACGCC	458
OY	138	GlyGluValGlnAlaMetLeuGluYglnSerThrGlnGluLeuAspValaArgLeuAlaSer	157
Db	459	AGCGAGGTGCACACATGTTGGCGACACACCGAGGAGGTGGGAGCGCCGTGGCTTCC	518
OY	158	HisteArGlySleuArGlyArgLeuLeuAspAlaAspAlaAspLeuGlnTyrArgLeu	177
Db	519	CACCTGGCCAAAGTCGGCAACGGGCTCTCCGCACACCGAGGACCTGCGAAGCGCTG	578
OY	178	AlaValTyrGlnAlaGlyAlaArgGluGluYAlaGluArgGlyLeuSerAlaIleArgGlu	197
Db	579	GCCGTATCCAGAGGGGGGCTGCGGAGGGCGCCGAGGCGACCGTGAACGGCTTCCGGAG	638
OY	198	ArgLeuGlyProLeuValGluGlnGlyArgValaArgAlaIaIaIaIaValGlySerLeuAla	217
Db	639	CGCCCTGGGCCCCGTGTGTGACAGAGGGCGATGTGGCGCCGCCACCTCGATACCGAGGCC	698
OY	218	GlyGlnProLeuGlnGluArgAlaGlnAlaTyrGlnGluArgLeuArgAlaTyrMetGlu	237
Db	699	GGCCAGCGCGTGGCGGAGCGCGGGAAGCCTGGGGGCCAGAACTGGCGGAGACGGCTGGAG	758
OY	238	GluMetGlySerArgThrArgAspArgLeuAspGluValYlsGluGlnValaIaVal	257
Db	759	GAGATGGGCACCCGAGCCCGCGACCGCGCTGATGATGATGCTGTGAGCGACGTGAGGAGGTG	818
OY	258	ArgAlaIysLeuGlnGluGlnAlaGlnGlnIleLeuArgLeuGlnAlaGlnAlaIaIa	277
Db	819	CGCACAAAGTGTGAGAGCACGAGCGACAGTGTGGCTGTGAGGCGCGGAGATTCCAGGCC	878

RESULT 4
US-08-726-306A-28
Sequence 28, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Butbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEO ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 base pairs

```

;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;      MOLECULE TYPE: cdna
US-08-726-306A-28

```

Pred. No.:	8,14e-88	Length:	660
Score:	984.00	Matches:	197
Percent Similarity:	98.99%	Conservative:	0
Best Local Similarity:	98.99%	Mismatches:	2
Query Match:	71.62%	Indels:	0
DB:	2	Gaps:	0

QY	1	MeLysValLeuTPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIysVal	20
Db	62	ATGAAGTCTCTGTGGGCTGCCTTCTGGTCAATCTCTGGAGAGATCCAGGCCAAGGTG	122
QY	21	GlulGlnAlaValAlaGluThrGluProGluIleuArgGlnGlnThrGluTrpGlnSer	40
Db	122	GAGCAAGCGGTGGAGACAGACCAGGACCCGAGCTGGCCAGCAGACGAGTGGCAGAGC	181
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyIleArgPheTrpAspIleuArgTrpValGlnThr	60
Db	182	GGCCAGCGCTGGGAACGGCCACTGGGTCGCTTTGGGATTAACCTGCCTGGTGCAACA	241
QY	61	LeuSerGluGlnValGlnGluIleuLeuSerSerGlnAlaThrGlnGluLeuArgAla	80
Db	242	CTGTCTAGACAGGTGCAGAGAGACTGTCTACGCTCCCAAGTACCCAGAACTAGAGGCG	301
QY	81	LeuMetAspGluThrMetIysGluLeuLysAlaThrIlySerSerGluLeuGluGlnLeu	100
Db	302	CTGATGGACGAGACCATTGAAGATTGAAGCCCTACAAATCGGAACCTGGAGAAACAATC	361
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerIySclLeuGlnAlaAlaGlnAla	120
Db	362	ACCCGGTAGCGGAGAGAGACCGCGGCACGCTGTCCAAAGAGGCTGCAGAGCGCGCAGGCC	421
QY	121	ArgLeuGluAlaAspMetGluAspValCysGlyIleArgLeuValGlnIlyArgGlyIleVal	140
Db	422	CGGTCGGGCGCGGACATGGAGAGACTGTGGCGCCCTGTGTGCAGTACCCGGCGAGGTG	481
QY	141	GlnAlaMetLeuGlyGlnSerThrGluGluIleuArgValArgLeuAlaSerHisLeuArg	160
Db	482	CAGGCCATAGCTCGGCGCAGAGACCGAGAGACTGGGGGTGGCGCTGCCACTGGCGC	541
QY	161	LysLeuArgIysArgLeuLeuArgAspAlaAspAspLeuGlnIlySarArgLeuAlaValTyr	180
Db	542	AAGCTGCCTAAGCGGCTCTCCCGGATCCCGATACCTGCAGAAAGCGCTGCAGAGTGTAC	601
QY	181	GlnIleAlaGlyAlaArgGlyIleAlaGluArgGlyIleuSerAlaIleArgGluIleArgLeu	199
Db	602	CAGGCGGGGGCGCGAGAGGGCGCGAGAGCGGCTTCAGCGCATTCGCGAGCGCTG	658

RESULT 5
US-08-949-155-51
: Sequence 51, Application US/08949155
: Patent No. 6271436
: GENERAL INFORMATION:
: APPLICANT: Piedraalta, Jorge A
: APPLICANT: Bazer, Fuller W
: TITLE OF INVENTION: Compositions and Methods for the
: TITLE OF INVENTION: Generation of Transgenic Animal Species
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ARNOLD, WHITE AND DURKEE
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: TX
: COUNTRY: US
: ZIP: 77210-4433

;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/819,964
;; FILING DATE: 28-Mar-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/949,155
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US 60/046,094
;; FILING DATE: 09-MAY-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hildier, David W.
;; REGISTRATION NUMBER: 41,071
;; REFERENCE/DOCKET NUMBER: TAMK:177
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (512) 418-3000
;; TELEFAX: (713) 789-2679
;; INFORMATION FOR SEQ ID NO: 51:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4267 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-819-964-51
;
Alignment Scores:
Pred. No.: 3,08e-75 Length: 4267
Score: 866.50 Matches: 192
Percent Similarity: 56.05% Conservative: 35
Best Local Similarity: 47.41% Mismatches: 40
Query Match: 63.06% Indels: 139
Gaps: 4
;
US-09-827-854-15_COPY_1_277 (1-277) x US-09-819-964-51 (1-4267)
QY 1 MetLysValIleuTrpAlaIleuValuThrPheLeuAlaGlySerGlnAlaLysVal 20
Db 2448 ATGGCTGTGAACCTGCTCGCT-----GCAGATGCGGACAGAGAGAC 2489
QY 21 GlnGlnAlaValaGlnThrGluProGluProGluLeuArg-----GlnGlnThrGlu 37
Db 2490 GAGCGCGGG-----CCGCGCGCGGAGGTGCACGTGCTGGGAGAGAGCCAG 2537
QY 38 TrpGlnSerGlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrp 57
Db 2538 TGGCAGGAGCAGCCAGCCCTGGGAGCAGGCCCTGGGCGCTTGGGATTACCTGGCGCTGG 2597
QY 58 ValGlnThrLeuSerGlnGlnValaGlnGlnGluLeuLeuSerSerGlnValaThrGlnGlu 77
Db 2598 GTGCAGTCCCTGTGTACCAAGTGCAGAGAGAGCTGTCAAGCACAAGTCAACCCAGAA 2657
QY 78 Leu----- 78
Db 2658 CT-GACGTAAAGTCCACCCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCTGA 2716
QY 78 ----- 78
Db 2717 CCTCCTGGCAACCGTGTCTTGAGACCTCAGGCTCAACCGCGGTTTCTTTCTG 2776
QY 78 ----- 78
Db 2777 TCCCTGTGCCAACCTCTGGGGGTCTGGGTCTCTGTTCTTTTTCCTTCCTCCTTTT 2836
QY 78 ----- 78
Db 2837 TTGGGGGGAAAAAACTTTTCTTTTCTTTTCTTTTCTTTGACTTCAATGCTTCTTTTCC 2896
QY 78 ----- 78
Db 2897 AATCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2956
QY 78 ----- 78

Db 2957 AATCTGCGACGCTCCTGGCCATCGCCAGCTCAGAGACCCCTCTTCCCTCACCGCC 3016
QY 79 -----ArgAlaLeuMetAspGluThrMetLysGluLeuLysAlaTyr 92
Db 3017 CCGGCCCTCTCTGCGCCAGGAGCTGTATAGAGAGAGATGAAGAGAGGTGAAGGCTTAC 3076
QY 93 LysSerGlnLeuGlnGlnGlnLeuThrProValAlaGlnGlnGlnGlnGlnGlnGlnGln 112
Db 3077 CCGAGAGACTGTGAGCGAGCGAGCTGGGCGCCCTGTACCCAGAGAGAGCGGCGCTGTC 3136
QY 113 LysGlnLeuGlnAlaAlaGlnAlaAlaArgLeuGlyAlaAspMetGlnLysGlyArg 132
Db 3137 AAGAGACTCAGAGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3196
QY 133 LeuValaGlnTyrArgGlyGlnValaGlnAlaMetLeuGlnSerThrGlnGlnLeuArg 152
Db 3197 TTGGTCTCTACCGCAGGAGGTGCACACATGTTGGCCAGACCCAGAGAGGTGCGG 3256
QY 153 ValArgLeuAlaSerHisLeuArgLysLeuArgLysLeuArgLysLeuArgLysLeuArg 172
Db 3257 AGCGGCTGCTGCTCCACCTGCGCAGAGTGCACAGCGGCGGCGGCGGCGGCGGCGGAG 3316
QY 173 LeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlnGlyValaGlnGlnGln 192
Db 3317 CTGCAGAACCGCTGCGCTGTACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3376
QY 193 SerAlaIleArgGlnArgLeuGlnProLeuValaGlnGlnGlyArgValaArgAlaThr 212
Db 3377 AGCGGCTGCGGAGGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3436
QY 213 ValGlySerLeuAlaGlyGlnProLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 232
Db 3437 CTGAGTACCAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAG 3496
QY 233 ArgAlaIleArgMetGlnGlnMetGlySerArgTrpArgAspArgLeuAspGluLysGlu 252
Db 3497 CCGGCGGCGCTGAGAGAGTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 3556
QY 253 GlnValaIleGlnValaArgAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 272
Db 3557 CAGCTGAGAGAGGTGGCGACCAAGTGAAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 3616
QY 273 GlnAlaPheGlnAla 277
Db 3617 GAGGAGTTCACGCGC 3631
;
RESULT 7
US-07-849-389-6
; Sequence 6, Application US/07849389
; Patent No. 5525493
; GENERAL INFORMATION:
; APPLICANT: HORNES, Erik
; APPLICANT: UHLEN, Mathias
; TITLE OF INVENTION: CLONING METHOD AND KIT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/849,389
; FILING DATE: 19920519
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16787/168/DFBC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)836-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 330 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: HUMAN LIPOPROTEIN E GENE
FEATURE:
NAME/KEY: CDS
LOCATION: 1..330
US-07-849-389-6

Alignment Scores:
Pred. No.: 5,71e-42 Length: 330
Score: 511.00 Matches: 107
Percent Similarity: 98.18% Conservative: 1
Best Local Similarity: 97.27% Mismatches: 2
Query Match: 37.19% Indels: 0
DB: 1 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-07-849-389-6 (1-330)

OY 87 LysGluLeuLysAlaTyrLysSerGluLeuGluGluLeuThrProValAlaGluGlu 106
DB 1 AAGGATTGAAGGCTCAGCAATCGAATCGAAGGAGGAACTACACCCGGTGGCGAGAG 60
OY 107 ThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGluAlaSpMet 126
DB 61 AGCGGGGCGAGGCTGTCCAAAGAGCTGACGCGCGGAGAGCCCGCTGGCGCGCAGT 120
OY 127 GluArgValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGluGln 146
DB 121 GAGGAGCTCGCGGCGCGCTGTGACAGTACCGCGGAGAGTGCAGGCCCATGCTCGGCCAG 180
OY 147 SerThrGluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeu 166
DB 181 AGCAGCGAGAGCTCGGGGTGGCGCTCCGCTCCAGCTCGCAAGCTGGTAAAGCGGCTC 240
OY 167 LeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGlu 186
DB 241 CTCGGGATGCCGATGACCTGCAGAACCGCTGGCAGTGTACAGAGCGCGGGCCGCGAG 300
OY 187 GlyAlaGluArgGlyLeuSerAlaIleArg 196
DB 301 GGCGCGAGCGCGCTCAGCGCCATCCGC 330

RESULT 8

US-08-617-256-24
Sequence 24, Application US/08617256
Patent No. 6043031

GENERAL INFORMATION:
APPLICANT: Kyster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,256
FILING DATE: March 18, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: March 17, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth A.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: SOI-013CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-617-256-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 3 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-08-617-256-24 (1-252)

OY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaArgLeuGlyAlaAspMetGluSp 128
DB 2 GCAGGCTGTCCAGAGAGCTGACGCGGCGGAGGCCCGCTGGCGGAGCATGAGAGAG 61
OY 129 ValCysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyInsThr 148
DB 62 GTGTGCGC-CGCTGTGTGACAGTACCGCGGAGGTGCAGGCCCATCTCGGCCAGACAC 120
OY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTCGGGGTGGCGCTCCGCTCCAGCTCGCAAGCTGGTAAAGCGGCTCCTCCGC 180
OY 169 AspAlaAspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluAla 188
DB 181 GATGCCGATGACCTGCAGAACGCTGGCAGTGTACAGAGCGCGGGCCGCGAGGCGGCC 240
OY 189 GluArgGlyLeu 192
DB 241 GAGCGCGGCTC 252

RESULT 9

US-09-287-141-24
Sequence 24, Application US/09287141
Patent No. 6197498

GENERAL INFORMATION:
APPLICANT: Kyster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Heller Ehrman White & Maulliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/287,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-587-5360
TELEFAX: 619-450-8400
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-287-141-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-141-24 (1-252)
QY 109 AlArGleuSerlysgluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCCGCGCTGGGGCGGCATGAGGAC 61
QY 129 ValCysGlyARgLeuValGlnTYrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGGC-CGCTGTGGTGCAGTACCGGGGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
Db 121 GAGGAGCTGGGGTGGCCCTCCGCTCCACCTGCGCAAGCTGGTAAAGGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyArgLeuAlaValTYrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTTCGAGAGTCTCTGGCAGTGTACAGGCCGGGGCCCGCAGGGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCCTC 252

RESULT 10
US-09-431-613-24
Sequence 24, Application US/09431613
Patent No. 6221601
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heller Ehrman White & McAuliffe
STREET: 4250 Executive Square, 7th Floor
CITY: La Jolla
STATE: CA

COUNTRY: USA
ZIP: 92037-9103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/431,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/617,256
FILING DATE: 18-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/406,199
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 24736-2002G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-587-5360
TELEFAX: 619-450-8400
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-431-613-24

Alignment Scores:
Pred. No.: 9.9e-29 Length: 252
Score: 374.00 Matches: 82
Percent Similarity: 97.62% Conservative: 0
Best Local Similarity: 97.62% Mismatches: 2
Query Match: 27.22% Indels: 1
DB: 4 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-431-613-24 (1-252)
QY 109 AlArGleuSerlysgluLeuGlnAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
Db 2 GCACGGCTGTCCAAAGAGCTGCAGCGCGCCGCGCTGGGGCGGCATGAGGAC 61
QY 129 ValCysGlyARgLeuValGlnTYrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
Db 62 GTGTGGGC-CGCTGTGGTGCAGTACCGGGGAGGTGCAGGCCATGCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
Db 121 GAGGAGCTGGGGTGGCCCTCCGCTCCACCTGCGCAAGCTGGTAAAGGGCTCTCCGC 180
QY 169 AspAlaAspAspLeuGlnTyArgLeuAlaValTYrGlnAlaGlyAlaArgGluGlyAla 188
Db 181 GATGCCGATGACTTCGAGAGTCTCTGGCAGTGTACAGGCCGGGGCCCGCAGGGCGCC 240
QY 189 GluArgGlyLeu 192
Db 241 GAGCGCGGCCTC 252

RESULT 11
US-09-504-245-24
Sequence 24, Application US/09504245
Patent No. 6221605
GENERAL INFORMATION:
APPLICANT: K ster, Hubert
TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry

DB 2 GCACGGCTGTCACAGAGCTGCAGCGCGCGCTGCGCGGACATGAGGAC 61
QY 129 ValcysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTACAGTACCGCGGAGGTGCAGGCCATCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTGCGCTCCACCTCGCGCAAGCTGCTGAAGCGGCTCTCCGC 180
QY 169 AspaIaaspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTACAGCGCGGCGCGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGCGCTC 252
RESULT 15
US-09-287-681-24
; Sequence 24, Application US/09287681
; Patent No. 6277573
; GENERAL INFORMATION:
; APPLICANT: K ster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McCauliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037-9103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,681
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-587-5360
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-287-681-24

Alignment Scores:
Pred. No.: 9.9e-29
Score: 374.00
Percent Similarity: 97.62%
Best Local Similarity: 97.62%
Query Match: 27.22%

Length: 252
Matches: 82
Conservative: 0
Mismatch: 2
Indels: 1

DB: 4 Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x US-09-287-681-24 (1-252)
QY 109 AlaArgLeuSerLysGluLeuGlnAlaAlaGlnAlaArgLeuGlyAlaAspMetGluAsp 128
DB 2 GCACGGCTGTCACAGAGCTGCAGCGCGCGCTGCGCGGACATGAGGAC 61
QY 129 ValcysGlyArgLeuValGlnTyrArgGlyGluValGlnAlaMetLeuGlyGlnSerThr 148
DB 62 GTGTCCGC-CGCTGTGTACAGTACCGCGGAGGTGCAGGCCATCTCGGCCAGAGCAC 120
QY 149 GluGluLeuArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArg 168
DB 121 GAGGAGCTGCGGGGTGCGCTGCGCTCCACCTCGCGCAAGCTGCTGAAGCGGCTCTCCGC 180
QY 169 AspaIaaspAspLeuGlnLysArgLeuAlaValTyrGlnAlaGlyAlaArgGluGlyAla 188
DB 181 GATGCCGATGACCTGCAGAGTCCCTGGCAGTACAGCGCGGCGCGGAGGCGCC 240
QY 189 GluArgGlyLeu 192
DB 241 GAGCGGCGCTC 252

Search completed: March 14, 2003, 20:18:31
Job time : 34.3514 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2003, 17:32:58 ; Search time 57.563 Seconds
(without alignments)
3377.460 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277
Sequence: 1 MKYLMALVTFLAGCAKY.....RAKLEQAQRIQLQAEAFQA 277

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 501302 seqs, 350932545 residues
Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09827854/rnat_11032003_101612_27578/app_query.fasta_1.3576
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsun62
-TRANS=Numan40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=US09827854.ecgn_1_1_298_@rnat_11032003_101612_27578
-MAXLEN=200000000 -USER=US09827854.ecgn_1_1_298_@rnat_11032003_101612_27578
-NCPU=6 -ICPU=3 -NO_XLIPX -NO_MAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/p2nmodel/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/p2nmodel/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/p2nmodel/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/p2nmodel/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/p2nmodel/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/p2nmodel/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/p2nmodel/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/p2nmodel/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/p2nmodel/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/p2nmodel/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/p2nmodel/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/p2nmodel/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/p2nmodel/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/p2nmodel/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1374	100.0	1156	9 US-09-870-759-129	Sequence 129, App
2	1374	100.0	1156	9 US-09-802-640-17	Sequence 17, Appl
3	1374	100.0	1156	10 US-09-827-854-8	Sequence 8, Appl1
4	1374	100.0	1291	12 US-10-044-090-454	Sequence 454, App

5	1370	99.7	1156	10 US-09-827-854-12	Sequence 12, Appl
6	1366	99.4	1156	10 US-09-827-854-9	Sequence 9, Appl1
7	1366	99.4	1156	10 US-09-827-854-11	Sequence 11, Appl1
8	1365	99.3	1157	10 US-09-954-456-760	Sequence 760, App
9	1365	99.3	1157	10 US-09-880-107-2244	Sequence 2244, App
10	1362	99.1	1156	10 US-09-827-854-7	Sequence 7, Appl1
11	1359	98.9	1156	10 US-09-827-854-10	Sequence 10, Appl1
12	1158.5	84.3	1156	10 US-09-967-013-5	Sequence 5, Appl1
13	656.5	47.8	786	10 US-09-925-302-133	Sequence 133, App
14	625	45.5	478	10 US-09-864-824-374	Sequence 374, App
15	625	45.5	478	10 US-09-880-107-2491	Sequence 2491, App
16	463	33.7	336	10 US-09-960-352-5420	Sequence 5420, App
17	458.5	33.4	449	10 US-09-960-352-4726	Sequence 4726, App
18	439.5	32.0	442	10 US-09-960-352-9395	Sequence 9395, App
19	435	31.7	414	10 US-09-960-352-4237	Sequence 4237, App
20	428.5	31.2	416	10 US-09-960-352-8720	Sequence 8720, App
21	422.5	30.7	423	10 US-09-960-352-14047	Sequence 14047, App
22	405.5	29.5	409	10 US-09-960-352-5148	Sequence 5148, App
23	401.5	29.2	416	10 US-09-960-352-9797	Sequence 9797, App
24	401	29.2	253	10 US-09-179-5368-130	Sequence 130, App
25	387.5	28.2	425	10 US-09-960-352-3497	Sequence 3497, App
26	380.5	27.7	390	10 US-09-960-352-1311	Sequence 1311, App
27	379.5	27.6	391	10 US-09-960-352-1278	Sequence 1278, App
28	379.5	27.6	392	10 US-09-960-352-10599	Sequence 10599, App
29	379.5	27.6	401	10 US-09-960-352-4187	Sequence 4187, App
30	375.5	27.3	413	10 US-09-960-352-8042	Sequence 8042, App
31	374	27.2	252	10 US-09-796-416-24	Sequence 24, Appl
32	374	27.2	252	10 US-09-879-341-24	Sequence 24, Appl
33	369.5	26.9	404	10 US-09-960-352-2325	Sequence 2325, App
34	369.5	26.9	405	10 US-09-960-352-14063	Sequence 14063, App
35	351.5	25.6	377	10 US-09-960-352-176	Sequence 176, App
36	350.5	25.5	353	10 US-09-960-352-4914	Sequence 4914, App
37	350.5	25.5	378	10 US-09-960-352-10040	Sequence 10040, App
38	350.5	25.5	366	10 US-09-960-352-10361	Sequence 10361, App
39	348.5	25.4	388	10 US-09-960-352-11886	Sequence 11886, App
40	344.5	25.1	396	10 US-09-960-352-3457	Sequence 3457, App
41	344.5	25.1	377	10 US-09-960-352-10540	Sequence 10540, App
42	338.5	24.6	432	10 US-09-960-352-8934	Sequence 8934, App
43	338	24.6	332	10 US-09-960-352-6552	Sequence 6552, App
44	337.5	24.6	376	10 US-09-960-352-5408	Sequence 5408, App
45	335.5	24.4	373	10 US-09-960-352-3681	Sequence 3681, App

ALIGNMENTS

RESULT 1
US-09-870-759-129
Sequence 129, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OR INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870, 759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 129
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (61)..(1014)
OTHER INFORMATION:
US-09-870-759-129

Alignment Scores:
Pred. No.: 1.01e-120
Score: 1374.00
Length: 1156
Matches: 277
Percent Similarity: 100.00%
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-870-759-129 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIysVal 20
 |||||||
 DB 61 ATGAAGGTTCTGTGGCTCGTTGCTGTCACATTCCTGGCAGGATGCCAGGCCAAGGTG 120
 QY 21 GluGlnAlaValAlaGluThrLupProGluPheLeuArgGlnThrGluTrpGlnSer 40
 |||||||
 DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCCGAGCTGCCAGCAGACCGAGTGGCAGAC 180
 QY 41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 DB 181 GGCACGCGCTGGGAAGTGGCTGCTTTGGGATTACTGGCGCTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 DB 241 CTGTCTGACAGGTGAGAGAGAGCTGCTCAGCTCCAGAGTCAACCAGAGACTGAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 |||||||
 DB 301 CTGATGAGAGAGACCATGAAGAGTTGAAGCCCTACAAATCGGAATCGAGGAGCAACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||||
 DB 361 ACCCGCGTGGCGAGAGAGACGGCGGCGACGCTGTCCAGAGAGCTGCAGCGCGCGAGCG 420
 QY 121 ArgLeuGlyAlaAspMetLysPvalCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 DB 421 CGGCTGGGCGGACATGAGAGACGTGTGCGCGCTGCTGTCAGATCCGCGCGAGGTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 481 CAGGCCATCTGCGGACAGACACCGAGAGCTGCGGCTGCTGCTCCACCTGGCC 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 541 AAGCTGCGTAAACGGGCTCTCCGCGATGCCGATGACCTCCAGAGCGCTGGCAGAGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 DB 601 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAGCCATCCGCGAGCGCTGGCG 660
 QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 DB 661 CCCCTGTGTGAACAGGCGCGGCGGCGGCGGCGGCTGCTGCTGCGCGCGCGCG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaArgMetGluGluMetGly 240
 |||||||
 DB 721 CTACAGAGAGCGGGCCAGGCGCTGGGCGGAGCGGCTGCGCGGATGAGAGAGATGGCG 780
 QY 241 SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArgAlaIys 260
 |||||||
 DB 781 ACCCGGACCCGCGACCGCTGTGACGAGGTGAAGAGAGCAGGTGGCGAGCTGGCGCGCAAG 840
 QY 261 LeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 |||||||
 DB 841 CTGGAGAGAGCGGCCAGCATAGCCTGCGAGCGCGAGCGCTTCAGAGCGC 891

RESULT 2

US-09-802-640-17

Sequence 17, Application US/09802640

Publication NO. US20030036057A1

GENERAL INFORMATION:

APPLICANT: Braun, Andreas

APPLICANT: Bonsai Aruna

APPLICANT: Kiehn Patrick

TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE

CURRENT APPLICATION NUMBER: US/09/802,640
 CURRENT FILING DATE: 2001-03-09
 NUMBER OF SEQ ID NOS: 122
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17

LENGTH: 1156

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: CDS

LOCATION: (61)...(1014)

OTHER INFORMATION: Nucleotide Sequence encoding apolipoprotein E

OTHER INFORMATION: (APOE)

US-09-802-640-17

Alignment Scores:

Pred. No.: 1,016-120

Score: 1374.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 9 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-802-640-17 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaIaIysVal 20
 |||||||
 DB 61 ATGAAGGTTCTGTGGCTCGTTGCTGTCACATTCCTGGCAGGATGCCAGGCCAAGGTG 120
 QY 21 GluGlnAlaValAlaGluThrLupProGluPheLeuArgGlnThrGluTrpGlnSer 40
 |||||||
 DB 121 GAGCAAGCGGTGGAGACAGAGCCGAGCCCGAGCTGCCAGCAGACCGAGTGGCAGAC 180
 QY 41 GlyLnaArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||||
 DB 181 GGCACGCGCTGGGAAGTGGCTGCTTTGGGATTACTGGCGCTGGGTGGCAGACA 240
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||||
 DB 241 CTGTCTGACAGGTGAGAGAGCTGCTCAGCTCCAGAGTCAACCAGAGACTGAGGGCG 300
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 |||||||
 DB 301 CTGATGAGAGAGACCATGAAGAGTTGAAGCCCTACAAATCGGAATCGAGGAGCAACTG 360
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||||
 DB 361 ACCCGCGTGGCGAGAGAGACGGCGGCGACGCTGTCCAGAGACTGCAGCGCGCGAGCGC 420
 QY 121 ArgLeuGlyAlaAspMetLysPvalCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||||
 DB 421 CGGCTGGGCGGACATGAGAGACGTGTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 481 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAGCCATCCGCGAGCGCTGGCGCGCAAG 540
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||||
 DB 541 AAGCTGCGTAAACGGGCTCTCCGCGATGCCGATGACCTCCAGAGCGCTGGCAGAGTAC 600
 QY 181 GlnAlaGlyAlaArgGluGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||||
 DB 601 CAGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||||
 DB 661 CCCCTGTGTGAACAGGCGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGlyLysArgLeuArgAlaArgMetGluGluMetGly 240
 |||||||
 DB 721 CTACAGAGAGCGGGCCAGGCGCTGGGCGGAGCGGCTGCGCGCGGATGAGAGAGATGGCG 780

Db	366	CTGTCTGACGACAGGTGCAGGAGGAGCTGCTTCACGCTCCACAGTCCACCCAGAACTGAGGCGC	425
QY	81	LeuMeLAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGluLeu	100
Db	426	CTGATGTGACAGAGCCCATGGAAGGATTGAAGGCCCTACAAATCGGAACCTGGAGGAACAAC	485
QY	101	ThrProValAlaGluLutThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Db	486	ACCCCGGTGGCGAGAGACACCGCGCACGGCTCTCCAGAGAGCTGCAGCGCGGCGAGGCC	545
QY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal	140
Db	546	CGGCTGGCGCGACATGAGAGACGTGTGGCGCGCCCTGGTGCAGTACCGCGCGAGGTG	605
QY	141	GlnAlaMetLeuGlyLysIleSerThrGluGluLeuArgValAlaArgLeuAlaSerHisLeuArg	160
Db	606	CAGGCGATCTCTGGCCAGACACCGAGGAGACTGCGGGTGGCTCCCTCCACCTGCGC	665
QY	161	LysIleuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr	180
Db	666	AAAGCTGCGTGAAGGGCTCTCCGCGCATGCCGATGCATGCAGAAAGCGCTGGGCACTGTAC	725
QY	181	GlnAlaGlyAlaArgGluGluGluAlaArgGlyLysSerAlaIleArgGluArgGluGly	200
Db	726	CAGCGCGGGGCCCGCGACAGGCGCGCGAGCGCGGCTTACAGCGCCATCCGCGAGCGCTTGGG	785
QY	201	ProLeuValGluGluGlnLysArgValAlaArgAlaIleThrValGlySerLeuAlaGlyIlePro	220
Db	786	CCCCTGTGTGAACAGAGGCGCGGTGGCGCGGCACACTGTGGCTCCTTGGCGCGGCGACGCG	845
QY	221	LeuGlnGluArgAlaGlnAlaIleTrpGlyLutArgLeuArgAlaArgMetGluMetGly	240
Db	846	CTACAGAGACGGGCGCCAGGCTGTGGGCGAGCGGCTCGCGCGGATGGAGAGATGGGC	905
QY	241	SerArgThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValAlaArgAlaLys	260
Db	906	ACCCGAGACCCGCGAGCGCTGTGCAGCAGGTGAAGAGACAGGTGGCGAGGTGGCGGCCAG	965
QY	261	LeuGluGluGlnAlaGlnGlnIleLeuArgLeuGlnAlaGluAlaPheGlnAla	277
Db	966	CTGAGAGACAGAGGCCAGACAGATACGCTGCAGAGCGGAGGCTTTCAGAGCC	1016
RESULT 5			
US-09-827-854-12			
: Sequence 12, Application US/09827854			
: Patent No. US20020123093A1			
: GENERAL INFORMATION:			
: APPLICANT: Zannis, Vassilis			
: APPLICANT: Kyriacos, Kyriakos E.			
: TITLE OF INVENTION: Compounds and methods for lowering			
: FILE REFERENCE: 07180/004003			
: CURRENT APPLICATION NUMBER: US/09/827, 854			
: PRIOR FILING DATE: 2001-04-05			
: PRIOR APPLICATION NUMBER: US 09/679, 088			
: PRIOR FILING DATE: 2000-10-04			
: PRIOR APPLICATION NUMBER: US 09/544, 386			
: PRIOR FILING DATE: 2000-04-06			
: NUMBER OF SEQ ID NOS: 19			
: SOFTWARE: FastSeq for Windows Version 4.0			
: SEQ ID NO 12			
: LENGTH: 1156			
: TYPE: DNA			
: ORGANISM: Homo sapiens			
US-09-827-854-12			
Alignment Scores:			
Pred. No.: 2,39e-120			
Score: 1370.00			
Percent Similarity: 100.00%			
Best Local Similarity: 99.64%			
Query Match: 99.71%			
DB: 10			
Length: 1156			
Matches: 276			
Conservative: 1			
Mismatches: 0			
Indels: 0			
Gaps: 0			

```

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-12 (1-1156)

Oy      1  MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal  20
Db      61  ATGAAGGTTCTGTGGCTGGCTGGTCAATTCCTGGCAGGATGCCAGGCCAAAGGTG  120

Oy      21  GluGlnAlaValAlaGluTrpGluProGluProGluLeuArgGlnGlnTrpGluTrpLysI  40
Db      121  GAGCAGAGGGTGGAGACAGAGCCGGAGACCAGACTCCGCCAGACAGCCAGTGGCAGAGC  180

Oy      41  GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr  60
Db      181  GGCCAGGCGCTGGCAACTGGCCACTGGGTCCTTTGGGATTTACCTGGCGTGGCTGCACACA  240

Oy      61  LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla  80
Db      241  CTGTCTGAGCAGGGTGCAGGAGGAGCTGCTCAGCTCCAGTCCAGTCAACCAGAACTGAGGGCG  300

Oy      81  LeuLeuAspGluTrpPheLeuLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu  100
Db      301  CTGATGGACGAGACCATGAAGAGACTTGAAGCCTACAAATTCGGAACTGGAGGAACAACATG  360

Oy      101  ThrProValAlaGluGluTrpArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla  120
Db      361  ACCCGGGGGGAGGAGAGACGCGGCGACTCCAAAGAGACTGCAGAGCGGCGCAGAGCC  420

Oy      121  ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGluVal  140
Db      421  CGGCTGGGGCGCGGCAATGAGGAGACGTGTGGGGCGCCCTGTCACAGTACCGCGGCAGAGTG  480

Oy      141  GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg  160
Db      481  CAGGCCATTGCTCGGCCAGAGCACCGAGAGACTGTGGGTGCGCTCGCTCCACCTGGCC  540

Oy      161  LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr  180
Db      541  AAGCTGCGTACAGCGGCTCTCCCGCATGCCGATGACCTCAAAACGCGCTGGCAGTGTAC  600

Oy      181  GlnAlaGlyAlaArgGluGlnAlaGluArgGlyLeuSerPheIleArgGluArgLeuGly  200
Db      601  CAGCGCGGGGGCGCGAGGGCGCCAGACCGCGGCTCTAGGCGCATCCGACGAGCGCTGGGG  660

Oy      201  ProLeuValGluGlnGlnArgValArgValAlaAlaThrValGlySerLeuAlaGlyLysPro  220
Db      661  CCCCTGGTGAACAGGGGCCGCTGTGGGGCGCCACTGTGGCTCCCTGGCCGGCCAGGCG  720

Oy      221  LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArgMetGluLeuMetGly  240
Db      721  CTACAGAGAGGGGGCCAGGCTGGGGCGAGCGGCTCCGCCCGATGAGAGATGGTGGCG  780

Oy      241  SerArgThrArgAspArgLeuAspGluValLysGlnGlnValAlaGluValArgAlaLys  260
Db      781  AGCCGAGCCCGCGACCCCTGGAGCAGGGTGAAGAGACAGTGGCGGAGGTGGGCCCAAG  840

Oy      261  LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla  277
Db      841  CTGAGAGAGCAGCGCCAGCAGATACCCCTGTCAAGCGCGAAGCCTTCAGAGCC  891

RESULT 6
US-09-827-854-9
: Sequence 9, Application US/09827854
: Patent No. US20020123093A1
: GENERAL INFORMATION:
: APPLICANT: Zambis, Vassilis
: APPLICANT: Kyrieos, Kyriakos E.
: TITLE OR INVENTION: Compounds and methods for lowering
: TITLE OR INVENTION: cholesterol levels without inducing hypertriglyceridemia
: FILE REFERENCE: 07180/004003
: CURRENT APPLICATION NUMBER: US/09/827, 854
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: US 09/679, 088
: PRIOR FILING DATE: 2000-10-04

```

; PRIOR APPLICATION NUMBER: US 09/544,386
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 9
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-827-854-9

Alignment Scores:

Pred. No.:	5,69e-120	Length:	1156
Score:	1366.00	Matches:	276
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.42%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-9 (1-1156)

Oy 1 MettysValleuTPRALAaleuValThrpheleuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGSGCTGCGTGTGTCACATTCTGGCAGATGCCAGGCCAAGGTG 120
Oy 21 GlnGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCTGGCCAGACAGACCGAGTGGCAGAGC 180
Oy 41 GlyGlnArgTTPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db 181 GCCCAGCGCTGGAGACTGGCAGCTGGTGGCTTTGGATTACCTGCGCTGGGTGCAGACA 240
Oy 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGATCCAGACCACTGAGGGCG 300
Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
Db 301 CTGATGAGCAGAGACCATGAGAGTTGAAGGCTTCAATCGAATCGAGAGAACACTG 360
Oy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTTGGGAGAGAACCGGCGCAGCGCTGTCCAAGAGACTCCAGCGCGCGCAGGCC 420
Oy 121 ArgLeuGlnAlaAspMetGlnAspValaCysGlyArgLeuValaGlnTyrArgGlyVala 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 480
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGTCTGGCCAGAGCACCAGAGGCTGGCGGTGGCGCTGCCACCTGGCGC 540
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGCTAAGCGGCTCTCCGCGATGCCATGTGACCTGCACAAATGCTGCGCACTGTAC 600
Oy 181 GlnAlaGlyAlaArgGlnGlyAlaGlnArgGlyLeuSerAlaHisLeuArgGlnArgLeuGly 200
Db 601 CAGCGCGGCGCGCGAGAGCGCGCCAGCGCGCTCAGCGCCATCCGAGAGCGCTGGGG 660
Oy 201 ProLeuValaGluGlnGlyArgValaArgAlaAlaThrValaGlySerLeuAlaGlyGlnPro 220
Db 661 CCCCTGGTGAACAGAGCGCGGTGGCGCGCCACTGTGGCTCTCCCTGGCGCGCCAGCGC 720
Oy 221 LeuGlnGlnAlaArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGlnGlnMetGly 240
Db 721 CTACAGGAGCGCGCCAGAGCTGGGGCGAGCGGTGGCGCGCGGATGAGAGATGGCGC 780
Oy 241 SerArgThrArgAspArgLeuAspGlyValLysGlnGlnValaGlnValaArgAlaLys 260
Db 781 ACCCGGAGCGCGAGCGCTGGACAGAGTGAAGAGAGGTGGCGAGAGTGGCGCCAG 840
Oy 261 LeuGlnGlnGlnAlaGlnGlnLysArgLeuGlnAlaGlnAlaPheGlnAla 277

Db 841 CTGAGAGAGAGCGCCAGACAGATACGCTGCGAGCGCGGCTTCCAGAGCC 891

RESULT 7

US-09-827-854-11
; Sequence 11, Application US/09827854
; Patent No. US20020123093A1
; GENERAL INFORMATION:

APPLICANT: Zannis, Vassilius

TITLE OF INVENTION: Compounds and methods for lowering

FILE REFERENCE: 07180/004003

CURRENT APPLICATION NUMBER: US/09/827,854

CURRENT FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: US 09/679,088

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 09/544,386

PRIOR FILING DATE: 2000-04-06

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11

LENGTH: 1156

TYPE: DNA

ORGANISM: Homo sapiens

US-09-827-854-11

Alignment Scores:

Pred. No.:	5,69e-120	Length:	1156
Score:	1366.00	Matches:	276
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	1
Query Match:	99.42%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-11 (1-1156)

Oy 1 MettysValleuTPRALAaleuValThrpheleuAlaGlyCysGlnAlaLysVal 20
Db 61 ATGAAGGTTCTGTGSGCTGCGTGTGTCACATTCTGGCAGATGCCAGGCCAAGGTG 120
Oy 21 GlnGlnAlaValaGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
Db 121 GAGCAAGCGGTGAGACAGAGCCGAGCTGGCCAGACAGACCGAGTGGCAGAGC 180
Oy 41 GlyGlnArgTTPGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTyrValGlnThr 60
Db 181 GCCCAGCGCTGGAGACTGGCAGCTGGTGGCTTTGGATTACCTGCGCTGGGTGCAGACA 240
Oy 61 LeuSerGluGlnValaGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
Db 241 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCAGCTCCAGATCCAGACCACTGAGGGCG 300
Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluLeu 100
Db 301 CTGATGAGCAGAGACCATGAGAGTTGAAGGCTTCAATCGAATCGAGAGAACACTG 360
Oy 101 ThrProValAlaGluGlnThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
Db 361 ACCCGGTTGGGAGAGAACCGGCGCAGCGCTGTCCAAGAGACTCCAGCGCGCGCAGGCC 420
Oy 121 ArgLeuGlnAlaAspMetGlnAspValaCysGlyArgLeuValaGlnTyrArgGlyVala 140
Db 421 CGCGTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGTCACTACCGCGGCGAGGTG 480
Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValaArgLeuAlaSerHisLeuArg 160
Db 481 CAGGCCATGTCTGGCCAGAGCACCAGAGGCTGGCGGTGGCGCTGCCACCTGGCGC 540
Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
Db 541 AAGCTGTGAAGCGGCTCTCCGCGATGCCATGTGACCTGCACAAATGCTGCGCACTGTAC 600

[illegible]

RESULT 8
 US-09-954-456-760
 : Sequence 760, Application US/09954456
 : Patent No. US20020115057A1
 : GENERAL INFORMATION:
 : APPLICANT: Young, Paul
 : TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cancer

```

1 CURRENT APPLICATION NUMBER: US/09/954,456
2
3 CURRENT FILING DATE: 2001-09-18
4
5 PRIOR APPLICATION NUMBER: US/60/233,617
6
7 PRIOR FILING DATE: 2000-09-18
8
9 PRIOR APPLICATION NUMBER: US/60/234,052
10
11 PRIOR FILING DATE: 2000-09-20
12
13 PRIOR APPLICATION NUMBER: US/60/234,923
14
15 PRIOR FILING DATE: 2000-09-25
16
17 PRIOR APPLICATION NUMBER: US/60/235,134
18
19 PRIOR FILING DATE: 2000-09-25
20
21 PRIOR APPLICATION NUMBER: US/60/235,637
22
23 PRIOR FILING DATE: 2000-09-26
24
25 PRIOR APPLICATION NUMBER: US/60/235,638
26
27 PRIOR FILING DATE: 2000-09-26
28
29 PRIOR APPLICATION NUMBER: US/60/235,711
30
31 PRIOR FILING DATE: 2000-09-27
32
33 PRIOR APPLICATION NUMBER: US/60/235,720
34
35 PRIOR FILING DATE: 2000-09-27
36
37 PRIOR APPLICATION NUMBER: US/60/235,840
38
39 PRIOR FILING DATE: 2000-09-27
40
41 PRIOR APPLICATION NUMBER: US/60/235,863
42
43 PRIOR FILING DATE: 2000-09-27
44
45 NUMBER OF SEQ ID NOS: 2276
46
47 SOFTWARE: PatentIn version 3.0
48
49 SEQ ID NO: 760
50
51 LENGTH: 1157
52
53 TYPE: DNA
54
55 ORGANISM: Homo sapiens
56
57 US-09-954-456-760

```

Pred. No.:	7.08e-120	Length:	1157
Score:	1365.00	Matches:	275
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	2
Query Match:	99.34%	Indels:	0
DB:	10	Gaps:	0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-954-456-760 (1-1157)

Oy 1 MethysValIeuTirPaIaAlaleuLeuValThrPheLeuAIGlyCysGlnAlaLysVal 20
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 ATGAAGCTTCGTGGGCTGCCGTTGCTGCACATTCCATGCGAGATGCCAGGCCAAGTG 120

OY	21	GIGlnAlaValAGluThrGluProGluProGluLeuMetArgGlnGlnThrGluUTPGLnSer	40
Dd	122	GAGCAGAGGGTGGACAGACCCGAGCCCGAATGCCCGCCAGCACAGCATGGCCAGACC	181
OY	41	GlyGlnArGTPrGluLeuAlaLeuGlyArGPhenTrpaSPyrLeuArGTPrPaLgInThr	60
Dd	182	GGCCAGGCCTGGAAACTGGCACTGGTGCTTTTGGATTACCTGGCGTGGGTGGACAGA	241
OY	61	LeuSerGIuGlnValGlnGluGluLeuLeuSerIserGlnValThyGlnGluLeuAlaGala	80
Dd	242	CTGTCTGGAGGTCAGAGAGACCTGCTACGCTCCCAAGTCACACCAAAGAACATGAGGGC	301
OY	81	LeuMetAPGJLuThrMetLysGluLeuLysAlaTyrrLysSergLueGluGluGluGlnLeu	100
Dd	302	CTGATGACGAGACCACTGAAGAGTTGAAAGCCCTCAAATTCGGAACTGGAGAACAACTG	361
OY	101	ThrProValAlaGluGluThrArgAlaArGTLeuSerLysGluLeuGlnAlaAlaGlnAla	120
Dd	362	A'CCCCGTAAGCGAGAGACCGGGSCACGGCTGTCCAAGAGGTGCAGACGGCGCAGGCC	421
OY	121	ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyraArgGlyGluVal	140
Dd	422	CGGCTGGCGCGGACATGGAAGACGTGTCGGCGCCCTGTGTCAGTACCCCGGCGAGGTG	481
OY	141	GlnAlaMetLeuGluGlnSerThrGlnGluLeuNryAlaArgLeuAlaSerHisLeuArg	160
Dd	482	CAGGCCATGCTGGCCCAAGCACCGAGAGAGTGGGGGTGGCCCTCGCTCCCACTGGCG	541
OY	161	LysLeuArGLysArGTLeuArGLeuArASpAlaAspAspLeuGlnLysArGTLeuAlaValTyr	180
Dd	542	AAGCTGCCTAAGCGGCTCTCGCGGATCCGATACCTCGCAAGAGCCCTGGCAGTGTAC	601
OY	181	GlnAlaGlyAlaArGTGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
Dd	602	CAGGCGGGGGCCCGCGAGGGCGCCGACGGCGCTTCAGCCCAATCCCGAGGCCCTGGGG	661
OY	201	ProLeuValIGluGlnGlyArgValArgAlaAlaThraGlySerLeuAlaGlyGlnPro	220
Dd	662	CCCCGGTGGGAACAGGGCGGTCGGGGCGCCCACTGTGGGCTCCCTGGCGGGCCAGCCG	721
OY	221	LeuGlnGluArGTAlaGlnAlaATPrGlyGluNryLeuArGTAlaArGTmetGluImetGly	240
Dd	722	CTACAGAGAGGGGCCAGGCCCTGGGGGAGGGCGTGGCGCGCGGAGTGAAGATGGGC	781
OY	241	SerArgThrArGaSPargLeuAspGluValLysGluGlnValAlaGluValArGTAlaLys	260
Dd	782	AGTCGAGACCCCGACCGGCTTGAGCAGAGTGAAGAGCAGGTGGCGGAGGTGGCGCCAG	841
OY	261	LeuGluGluGlnAlaGlnGlnIleArGTLeuGlnIleAlaGlnAlaPheGlnAla	277
Dd	842	CTGGAGGAGCAGGCCCAACAATACGCTTGAGGCGCCAGGCGCTTCCAGGGCC	892

RESULT 5
US-09-88

; Sequence 2244, Application US/09880107
; Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

```

; APPLICANT: Schert, Uwe
: APPLICANT: Gene Ionic, Inc.

```

APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,1

; CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US
PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 3950

```

; NUMBER OF SEQ ID NOS. 3550
;
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 2244

; LENGTH: 1157

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 M12529
US-09-880-107-2244

Alignment Scores:
Pred. No.: 7,08e-120 Length: 1157
Score: 1365.00 Matches: 275
Percent Similarity: 99.28% Conservative: 0
Best Local Similarity: 99.28% Mismatches: 2
Query Match: 99.34% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-880-107-2244 (1-1157)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 62 ATGAAGGTTCTGTGGCTGCTGCTGTGTCACATTCTGGCAGATGCGCAGGCAAGGTG 121
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnProGlnSer 40
DB 122 GAGCAAGCGGTGGAGACAGACGCCGAGCTGCCAGCAGACGACGAGCGCAGAGCC 181
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 182 GCGCAGCGCTGGAGACTGGCAGTGGCTTTGGATTAACCTGGCGGTGCGACACA 241
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 242 CTGTGTGACAGTCAGAGAGAGCTGCTCAGCTCCCAAGTCACCCAGAACCTGAGGGCG 301
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
DB 302 CTGATGAGAGACCATGAGAGGTGAGAGGCTTACAAATCGGAAGTGGAGGACAACTG 361
QY 101 ThrProValaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 362 ACCCGGTAGCGGAGAGAGCGGCGCAGCGCTGTCCAAAGACTGCGAGCGCGCAGCGC 421
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 422 CCGCTGGCGCGGACATGAGAGAGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 481
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnArgValArgLeuAlaSerHisLeuArg 160
DB 482 CAGGCCATGCTGGCCAGAGACACCGAGAGCTGGGGGTGGCTGCTGCTGCCACTGGCC 541
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 542 AAGCTGCGTAAGCGGCTCTCCGCGATCCGATGACCTGACCAAGCGCTGGCAGTGTAC 601
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 602 CAGGCGGGGGCGCGGAGGCGCGCAGCGCGCTCAGCGCCATCGCGAGCGCTGGGG 661
QY 201 ProLeuValGlnGlnGlyArgValArgValAlaThrValGlySerLeuAlaGlyGlnPro 220
DB 662 CCCCTGGTGGAGACAGGGCGCGTGGGGCGCGCTGTTGGCTGCTGGCGCGCAGCGG 721
QY 221 LeuGlnGluArgAlaGlnAlaThrPylGluArgLeuArgAlaArgMetGlnGlnMetGly 240
DB 722 CTACAGAGAGCGGCGCGCGCGCGCGCGCTGCGCGCGCGCGATGGAGAGATGGGCG 781
QY 241 SerArgThrArgAspArgLeuAspGlnValLysGlnGlnValAlaGluValArgAlaLys 260
DB 782 ACTCGAGACCGCGCGCGCTGCGAGCAGAGGAGAGAGGTGGCGAGGTGGCGCGCAAG 841
QY 261 LeuGlnGluGlnAlaGlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
DB 842 CTGGAGAGACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGGCGC 892
RESULT 10

US-09-827-854-7
Sequence 7, Application US/09827854
Patent No. US20020123093A1
GENERAL INFORMATION:
APPLICANT: Zannis, Vassilis
APPLICANT: Kyriacos E.
TITLE OF INVENTION: Compounds and methods for lowering
FILE REFERENCE: 07180/004003
CURRENT APPLICATION NUMBER: US/09/827,854
CURRENT FILING DATE: 2001-04-05
PRIOR APPLICATION NUMBER: US 09/679,088
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 09/544,386
PRIOR FILING DATE: 2000-04-06
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1156
TYPE: DNA
ORGANISM: Homo sapiens
US-09-827-854-7

Alignment Scores:
Pred. No.: 1.35e-119 Length: 1156
Score: 1362.00 Matches: 276
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 99.13% Indels: 0
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-827-854-7 (1-1156)

QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB 61 ATGAAGGTTCTGTGGCTGCTGCTGTGTCACATTCTGGCAGATGCGCAGGCAAGGTG 120
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGlnProGlnSer 40
DB 121 GAGCAAGCGGTGGAGACAGACGCCGAGCTGCCAGCAGACGACGAGCGCAGAGCC 180
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
DB 181 GCGCAGCGCTGGAGACTGGCAGTGGCTTTGGATTAACCTGGCGGTGCGACACA 240
QY 61 LeuSerGluGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 241 CTGTGTGACAGTCAGAGAGAGCTGCTCAGCTCCCAAGTCACCCAGAACCTGAGGGCG 300
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnGlnLeu 100
DB 301 CTGATGAGAGACCATGAGAGGTGAGAGGCTTACAAATCGGAAGTGGAGGACAACTG 360
QY 101 ThrProValaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
DB 361 ACCCGGTAGCGGAGAGAGCGGCGCAGCGCTGTCCAAAGACTGCGAGCGCGCAGCGC 420
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTyrArgGlyGlnVal 140
DB 421 CCGCTGGCGCGGACATGAGAGAGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 141 GlnAlaMetLeuGlyGlnSerThrGlnGlnArgValArgLeuAlaSerHisLeuArg 160
DB 481 CAGGCCATGCTGGCCAGAGACACCGAGAGCTGGGGGTGGCTGCTGCCACTGGCGC 540
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
DB 541 AAGCTGCGTAAGCGGCTCTCCGCGATGACCTGACCAAGCGCTGGCAGTGTAC 600
QY 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
DB 601 CAGCGCGGGCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCAGGCG 660
RESULT 10


```
|||||
Db 20436 TGGGATTACCTCGCTGGGTGACAGACTGCTGTGACAGAGTGCAGAGAGTGCCTCAGC 20495
Oy 72 SerGlnValThrGlnGluLeu-----78
Db 20496 TCCAGAGTACCCAGAGAACTGAGGTGAGTGCCTCCATCCCTGAGCCTTGAGCTCTGGTG 20555
Oy 78 -----78
Db 20556 GCGGCTATACCTCCCAAGTCCAGGTTTCATTCTGCCCTGTGCTAAGTCTTGGGGG 20615
Oy 78 -----78
Db 20616 CTTGGGCTCTGCTGCTGTAGCTTCTCTTCCATTTCTGACTCTGCTGCTTACCTCTC 20675
Oy 78 -----78
Db 20676 TGGAAATCTCTCTCTCAGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 20735
Oy 78 -----78
Db 20736 CTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20795
Oy 78 -----78
Db 20796 CTACTGTTGTTGCCAGGCTGTGCTTGAACCTTCTGGGCTCAAGGAGATCTCTCCGCTCG 20855
Oy 78 -----78
Db 20856 CCTCCAAAGTGTGGAGATTAGAGCCATGAGCCACTTGGCCGCTCTGACTCTCTCT 20915
Oy 78 -----78
Db 20916 TCGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 20975
Oy 78 -----78
Db 20976 GCCTGCGCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 21035
Oy 78 -----78
Db 21036 CCCCATCCAGCCCTTCTCCCGCTCCACCTGTCGACACCCCTCCCGCTCTGCGCG 21095
Oy 79 -ArgAlaLeuMetAspGluThrMetTyrGlnValSerGlnLeuGlu 98
Db 21096 CAGGCGCTGATGAGCAGACCATTAAGAGTTGAAGGCTTCAATATGGAACTGGAGGA 21155
Oy 98 uGlnLeuThrProValAlaGluGluThrArgAlaArgLeuSerTyrGlnLeuGlnAla 118
Db 21156 ACAACTGACCCCGTGGGAGAGAACCGGGCAGCGCTGTCCAGAGAGCTGCAGCGG 21215
Oy 118 aglnAlaArgLeuGlnAlaAspMetGluAspValCysGlnArgLeuValGlnTyrArg 138
Db 21216 GAGAGCCGCGGGGCGGAGATGAGAGCTGTGGCGCCCTGGTGCAGTACCGCGG 21275
Oy 138 yGlnValGlnAlaMetLeuGlnSerThrGlnGluLeuThrArgValArgLeuAlaSer 158
Db 21276 CGAGGTGAGGAGCATGCTCGGCGACAGACAGAGAGCTGCGGGTGGCCTCTCCCA 21335
Oy 158 sLeuArgLysLeuArgLysArgLeuLeuArgAspAlaAspSerLeuGlnTyrArg 178
Db 21336 CCTGGCAGAGCTGTAGAGCGCTCTCCGCGATGCCATGACCTGCAGAGAGCGCTGG 21395
Oy 178 aValTyrGlnAlaGlnAlaArgGlnGluAlaGlnArgLysLeuSerAlaLeuArgVal 198
Db 21396 AGTTGACCAAGGCGGGGCGGAGAGGCGCGAGCGCGCTCCACAGCGCATCCGGAGCG 21455
Oy 198 gLeuGlyProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAla 218
Db 21456 CTTGGGCGCCCTGTGTGAACAGGCGGCTGGGCGCCCACTGTGGGCTCTCTGGCGG 21515
Oy 218 yGlnProLeuGlnArgAlaGlnAlaTrpGlyGlnArgLeuArgAlaArgMetGlu 238
|||||
```

```
Db 21516 CCAGCCGTACAGAGAGCGGCCCAAGCCCTGGGCGAGCGGCTGCCCGCGGATGAGGA 21575
Oy 238 uMetGlySerArgThrArgAspArgLeuAspGluValLysGlnValAlaGlnValArg 258
Db 21576 GATGGGACAGCCGAGACCCCGAGCCGCTGAGAGGTGAAGAGCAGGTGGCGAGGTGCG 21635
Oy 258 gAlaLysLeuGlnGlnAlaGlnGlnAlaLeuArgLeuGlnAlaGlnAlaAla 277
Db 21636 CGCCAAAGCTGAGAGAGAGCGGCCACAGATACGCTGAGAGCGCGAGGCTTCCAGGCC 21693

RESULT 13
US-09-925-302-133
; Sequence 133, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 133
; LENGTH: 786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-133

Alignment Scores:
Pred. No.: 1,85e-53 Length: 786
Score: 656.50 Matches: 143
Percent Similarity: 97.95% Conservative: 0
Best Local Similarity: 97.95% Mismatches: 3
Query Match: 47.78% Indels: 2
DB: 10 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x US-09-925-302-133 (1-786)
Oy 132 ArgLeuValGlnTyrArgGlyGlnValGlnAlaMetLeuGlnSerThrGlnGluLeu 151
Db 6 CGCCTGTGTCAGTACCGGCGGAGGTGCASGCAATCTCGGCCAGACAGCAGAGAGCTG 65
Oy 152 ArgValArgLeuAlaSerHisLeuArgLysLeuArgLysArgLeuArgAspAlaAsp 171
Db 66 GGGTGGCCCTCGCTCCACCTGCGCAA-CTGCGTAA-CGGCTCTCCGCGATGCCAT 123
Oy 172 AspLeuGlnLysArgLeuAlaValTyrGlnAlaGlnValArgGlnGluValArgGly 191
Db 124 GACCTGCAGAAAGCGCTGCAATGACCAAGCGGCGCGGAGAGGCGCGAGCGGCG 183
Oy 192 LeuSerAlaLeuArgGlnArgLeuGlyProLeuValGlnGlnGlyArgValArgAla 211
Db 184 CTCAGCGCATCCGCGAGCGCTGGGCGCCCTGGTGAAGAGGCGCGGTGGCGCGCC 243
Oy 212 ThrValGlySerLeuAlaGlyGlnProLeuGlnGlnAlaGlnAlaTrpGlyGlnArg 231
Db 244 ACTGTGGCTCCCTGGCGCGCGAGCGCTACAGAGCGCGCGCGCTGGGCGAGCGG 303
Oy 232 LeuArgAlaArgMetGlnGlnMetGlySerArgThrArgAspArgLeuAspGluValLys 251
Db 304 CTGCGCGCGGAGTGAAGAGATGGGAGAGCGGAGCGCGCGAGCGCTGAGAGGTGAG 363
Oy 252 GlnGlnValAlaGlnValArgAlaLysLeuGlnGlnAlaGlnGlnAlaTrpGlnGln 271
Db 364 GAGCAGGTGGGAGAGGTGCGCGCAAGCTGGAGGAGCGCCAGACATACGCTCGAG 423
Oy 272 AlaGlnAlaPheGlnAla 277
Db 424 GCCGAGCGCTTCCAGGCC 441
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 14, 2003, 12:09:47 ; Search time 1126.96 Seconds

(without alignments)
3980.771 Million cell updates/sec

Title: US-09-827-854-15_COPY_1_277

Perfect score: 1374

Sequence: 1 MKYLMALLTVFLGCAQV.....RAKLEQAQIRLQAEAFQA 277

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlp

-O=/cgn2_1/USPFO.spool/US09827854/rnat_11032003_101610_27495/app_query.fasta_1.3576

-DB=EST -QMT=fastcat -SUFFIX=p2n.rst -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USRR=US09827854.@cgn.1.1.8826.@rnat.11032003_101610_27495 -NCPU=6 -ICPU=3

-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120

-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estha:*

2: em_esthum:*

3: em_estlin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hnc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hnc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estom:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vtl:*

22: em_gss_fun:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1254	91.3	842	12	BG763371
2	1225.5	89.2	933	13	B1668318
3	1212.5	88.2	1027	13	B1670367
4	1203	87.6	800	13	BM042094
5	1202.5	87.5	942	13	B1600906
6	1199.5	87.3	922	13	B1597743
7	1190.5	86.6	817	12	BG774871
8	1189	86.5	811	13	B1600553
9	1187.5	86.4	938	12	BG761746
10	1175.5	85.6	845	12	BG829472
11	1171	85.2	927	12	BG472299
12	1152	83.8	790	12	BG707147
13	1149	83.6	919	13	B1551475
14	1139	82.9	757	13	BM042228
15	1129.5	82.2	907	12	BG706129
16	1128	82.1	706	14	BM728696
17	1128	82.1	741	12	BG762924
18	1124	81.8	804	12	BG702752
19	1120.5	81.6	808	13	B1668329
20	1117	81.3	855	13	B161362
21	1114	81.1	803	13	B1670350
22	1111	80.9	782	12	BG716776
23	1111	80.9	812	13	B1601551
24	1111	80.9	965	14	BQ677266
25	1108	80.6	757	12	B1603658
26	1105.5	80.5	797	12	BG715366
27	1105	80.4	846	13	B1159757
28	1104	80.3	796	13	BM042153
29	1099	80.0	757	12	BG707750
30	1099	80.0	794	13	B1601279
31	1097	79.8	790	13	B1551066
32	1096	79.8	790	13	B1551811
33	1092.5	79.5	798	12	BG708414
34	1091.5	79.4	891	13	B1549292
35	1085.5	79.0	802	13	B1458355
36	1084	78.9	1100	14	BM914382
37	1083.5	78.9	914	13	B1603523
38	1083	78.8	748	13	B1533085
39	1078.5	78.5	812	12	BG769968
40	1077.5	78.4	803	12	BG709360
41	1075	78.2	845	13	B1160936
42	1071	77.9	821	13	B1666579
43	1068.5	77.8	903	13	B1552527
44	1067.5	77.7	810	12	BG714022
45	1067	77.7	757	12	BG703504

ALIGNMENTS

RESULT 1
LOCUS BG763371 842 bp mRNA linear EST 15-MAY-2001
DEFINITION 602735433F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4680585 5',
ACCESSION BG763371
VERSION BG763371.1 GI:14074024
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 842)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LNCM1718 row: b column: 10
 High quality sequence stop: 817.
 Location/Qualifiers

FEATURES

1..842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4860585"
 /clone_1lb="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 154 a 250 c 337 g 101 t
 ORIGIN

Alignment Scores:

Pred. No.: 4,62e-123 Length: 842
 Score: 1254.00 Matches: 260
 Percent Similarity: 97.04% Conservative: 2
 Best Local Similarity: 96.30% Mismatches: 7
 Query Match: 91.27% Gaps: 2
 DB: 12

US-09-827-854-15_COPY_1_277 (1-277) x BG763371 (1-842)

OY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||||
 DB 36 ATGAAGTTCTGTGGCTCGTGGTGTGACATTCCTGGAGAGATCCAGGCCAAGG 95
 OY 21 GlnGlnAlaValAlaGlnThrGlnProGlnProGlnLeuArgGlnGlnThrGlnIns 40
 |||||||
 DB 96 GAGCAAGCGGTGGAGACAGACCCGAGCGCTGCCACAGACCCAGTGGCGAGAC 155
 OY 41 GlyGlnArgTrpGlnLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGln 60
 |||||||
 DB 156 GGCACAGCGCTGGGAACTGGCTGCGCTTTGGATTACCTGGCTGGTGCAGACA 215
 OY 61 LeuSerGlnGlnValGlnGlnGlnLeuSerSerGlnValThrGlnGlnLeuArgAla 80
 |||||||
 DB 216 CTGTCTGACGACGTGAGGAGAGCTGCTCAGCTCCACAGTCAACCAAGTGAAGG 275
 OY 81 LeuMetAspGlnThrMetLysGlnLeuLysAlaTrpLysSerGlnLeuGlnGlnLeu 100
 |||||||
 DB 276 CTGATGAGACAGACCATGAGAGGCTTGAAGGCTTCAAAATCGGAACGTGAGACAAC 335
 OY 101 ThrProValAlaGlnGlnThrArgAlaArgLeuSerLysGlnLeuGlnAlaGlnAla 120
 |||||||
 DB 336 ACCCGGTGGCGGAGAGACCGCGGCGCTGTCCAAAGAGACTGAGCGCGCAGCGC 395
 OY 121 ArgLeuGlnAlaAspMetGlnAspValGlyArgGlnLeuAlaGlnThrArgGlnVal 140
 |||||||
 DB 336 CGGCTGGCGCGGACATGAGAGCGGTGCGCGCTGTGTACAGTACCGCGCAGG 455
 OY 141 GlnAlaMetLeuGlnGlnSerThrGlnGlnLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||||
 DB 456 CAGGCCATGCTCGGCGACAGACCCAGAGAGCTGCGGCTGCGCTGCCCTCCACCTGCC 515

OY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTrp 180
 |||||||
 DB 516 AAGCTGCTGAAGGGGCTCTCCCGCGATGACCTGCACAGAGCGCTGCGACGTAC 575
 OY 181 GlnAlaGlnAlaArgGlnGlnValAlaGlnValLeuSerAlaIleArgGlnArgLeuGly 200
 |||||||
 DB 576 CAGCGCGGGCGCGGAGGCGCGCGAGCGCGCTGAGCGCATCCGCGAGCGCTGGGG 635
 OY 201 ProLeuValGlnGlnGlnArgValArgAlaAlaThrValGlnSerLeuAlaGlnPro 220
 |||||||
 DB 636 CCTGTGTCGACAGAGGCGCGCGTGGCGCGGACAGTGTGCGCTGCGCGGCGCG 694
 OY 221 LeuGlnGlnArgAlaGlnAlaThrPheGlnGlnArgLeuArgAlaArgMetGlnMet 240
 |||||||
 DB 695 CTACAGAGACCGGCGCGCGCTGCGCGGCGAGCGCTGCGCGCGATGAGAGATGCGG 754
 OY 240 YserArgTrpArgAspArgLeuAspGlnValLysGlnGlnValAlaGlnValArgAla 260
 |||||||
 DB 755 CAGCGGACCGCGCGAGCGCTGACAGCAGTGAAGAGGACAGTGGCGAGGTGCGGAC 814
 OY 260 LysGlnGlnGlnAlaGlnGlnLeuArg 269
 |||||||
 DB 815 CTGGGAGAGAGACGCCACACATACGC 842

RESULT 2

B1668318 933 bp mRNA linear EST 12-SEP-2001
 B1668318 603295681p1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5314844 5',
 LOCUS
 DEFINITION
 mRNA sequence.

VERSION B1668318
 KEYWORDS B1668318.1 GI:15582551
 SOURCE EST.

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE 1 (bases 1 to 933)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LNCM1798 row: a column: 21
 High quality sequence stop: 780.
 Location/Qualifiers

FEATURES

source

1..933
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5314844"
 /clone_1lb="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."
 BASE COUNT 171 a 284 c 359 g 119 t
 ORIGIN

Alignment Scores:		5.74e-120	Length: 933
Pred. No.:	1225.50	Matches: 259	
Score:	94.64%	Conservative: 6	
Percent Similarity:	92.50%	Mismatches: 12	
Best Local Similarity:	89.19%	Indels: 4	
Query Match:	13	Gaps: 1	
DB:	13		
US-09-827-854-15_COPY_1_277 (1-277) x B1668318 (1-933)			
QY	1	MethylsValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal	20
DB	75	ATGAAGGTTCTGTGGCGCTGCTGTGTGTCACATTCTGGCAGATGCCAGGCAAGGTG	134
QY	21	GlunAlaValaGlunThcGluProGluProGluLeuArgGlnThcGluThrPgiInser	40
DB	135	GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGACAGACCAGAGTGGCAGAC	194
QY	41	GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr	60
DB	195	GGCCAGCGCTGGAGACTGGCACTGGCTTTGGGATTACTCGCGTGGTGCAGACA	254
QY	61	LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThcGlnGluLeuArgAla	80
DB	255	CTGTCTGACAGAGTCAGAGAGAGCTGCTCACCTCCAGGTCAACCAGAGAACTGAGGGCG	314
QY	81	LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu	100
DB	315	CTGATGGACGAGACCATGAGAGTTGAAAGCCCTACAAATCGGAATGGAGAGACAACTG	374
QY	101	ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla	120
DB	375	ACCCCGGTGGCGAGAGAGCGCGGACAGCGCTGCCAAGAGACTGACAGCGCGCAGGCG	434
QY	121	ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal	140
DB	435	CGGCTGGGGCGGACATGAGAGAGCTGTCCGCGCTGTGCTGACATCCGCGCGCAGGTG	494
QY	141	GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg	160
DB	495	CAGGCCATGCTGGCCAGAGACCCAGAGAGCTGGGGTGGCTGCTCCCACTGCGCC	554
QY	161	LysLeuArgLysArgLeuLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTrp	180
DB	555	AAGCTGCGTAAGCGGCTCTCCGCGCATGTGACCTGCAGAAAGCGCTGGCAGTGTAC	614
QY	181	GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly	200
DB	615	CAGCGCGGGCGCGGAGGGCGCCAGAGCGCGCTCAGCGCCATCCGCGAGGCGCTGGGG	674
QY	201	ProLeuValGluGlnGlyValArgValArgAlaIleThrValGlySerLeuAlaGlyGlnPro	220
DB	675	CCCCCTGGTGGAGACAGAGCGCGCTGGCGCGCCACTGTGGCTCCCTGGCCAGCGCG	734
QY	221	LeuGlnGluArgAlaGlnAlaIleTrpGlyGluArgLeuArgAlaArg-MetGlu---GluMet	239
DB	735	CTACAGAGAGCGCGCCAGGCGCTGGGGCAGCGCGCTGGCGCGCGGATTTGAGACCATTTG	794
QY	239	IleYserArgTrpArgAspArgLeuAsp-GluValLysGlnGlnValAlaGluValArgA	259
DB	795	GGGCAAGCGGTGGAGACCCGCTGGAAGAGAGAGAGAGAGTGGGAGAGTGGCGCC	854
QY	259	IalYsLeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGluAlaPheGlnAla	277
DB	855	GCCAAAGGCTGGGGGAGACAGGCCGCCCAATTCCGCTTCAGGCCAA-GCCTTCACAGGCC	909
RESULT 3			
LOCUS	B1670367	1027 bp	mrna linear EST 12-SEP-2001
DEFINITION	603292738F1 NIH_MGC_96	Homo sapiens	cdna clone IMAGE:5312024 5',
ACCESSION	B1670367		
VERSION	B1670367.1	GI:15584600	

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1027)
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bbs-remail.nih.gov Tissue Procurement: Miklos Palokivits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L14M11790 row: 1 column: 09 High quality sequence stop: 845.
FEATURES	Location/Qualifiers
source	1..1027 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5312024" /clone_lib="NIH_MGC_96" /tissue_type="hypothalamus" /lab_host="DH10B" /note="Organ: brain; Vector: pbluescript (modified pbluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag) ; Oligo-dt primed using primer 5'-tttttttttttttttVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT	194 a 308 c 400 g 125 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.59e-118 Length: 1027
Score:	1212.50 Matches: 263
Percent Similarity:	93.97% Conservative: 2
Best Local Similarity:	93.26% Mismatches: 10
Query Match:	88.25% Indels: 7
DB:	13 Gaps: 1
US-09-827-854-15_COPY_1_277 (1-277) x B1670367 (1-1027)	
QY	1 MethylsValLeuThrPAlaAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
DB	26 ATGAAGGTTCTGTGGCGCTGCTGTGTGTCACATTCTGGCAGATGCCAGGCAAGGTG 85
QY	21 GlunAlaValaGlunThcGluProGluProGluLeuArgGlnThcGluThrPgiInser 40
DB	86 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCCAGACAGACCAGAGTGGCAGAC 145
QY	41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB	146 GGCCAGCGCTGGAGACTGGCACTGGCTTTGGGATTACTCGCGTGGTGCAGACA 205
QY	61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThcGlnGluLeuArgAla 80
DB	206 CTGTCTGACAGAGTCAGAGAGAGCTGCTCACCTCCAGGTCAACCAGAGAACTGAGGGCG 265
QY	81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluGlnLeu 100
DB	266 CTGATGGACGAGACCATGAGAGTTGAAAGCCCTACAAATCGGAATGGAGAGAACTG 325
QY	101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120

Db 326 ACCCGGTGGCGGAGAGAGCGGCGCACGGCTGTCAAAGAGCTGCAGGCGCGCAGGCC 385
 QY ArgLeuGluYAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 386 CGGCTGGGCGCGGACATGAGAGAGCTGTGGCGCGCTGTGTGATACCCGCGGAGAGTG 445
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 446 CAGGCGCATGCTGGCGCCAGACACCGAGAGAGCTGGGGTGGCTCGGCTCCACACTGCGC 505
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 506 AACCTGGTAAAGCGGCTCCCGCGATGCGATGACCTCGAGAAGCCCTGGGAGAGTGTAC 565
 QY 181 GlnAlaGlyAlaArgGluGluGluValArgGlyLysSerAlaIleArgGluArgLeu-GI 200
 Db 566 CAGGCGCGGCGCGCGAGGCGCGCGAGCGGCGCTCAGCCCATCCGCGAGCGCGCTGGG 625
 QY 200 YProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPr 220
 Db 626 CCCCCTGTGGAAACAGGTCCGCGCGCGCGCCACTGTGGCTCCCTGGCGCGCAGCC 685
 QY 220 GluGlnGluArgAlaGlnAlaThrPglGluArg-LeuArgAlaArgMetGlu-Glu-Met 239
 Db 686 GCTACAGAGAGCGGCGCAAGCGCTGGGCGCGCGCTTGCAGCGCGGATGGAGAGAGG 745
 QY 240 GlySerArg-ThrArgAspArgLeuAspGluValLysGluGluValAlaGlu-ValArgA 259
 Db 746 GGGAGCGCGGAGCCCGGACCGCTGACCGAGGTGAAAGAGACAGTGGCGGAAAGTGGCG 805
 QY 259 Jalysleu---GluGluGlnAlaGlnGln-IleArgLeuGlnAlaGluAlaPhe 275
 Db 806 CCCAAGCTGGAGCGAGCGAGCCAGCAAGATACGCTCGCAGCGCGCGAGGCTTC 859
 RESULT 4
 LOCUS BM042094 800 bp mRNA linear EST 07-NOV-2001
 DEFINITION 603615713F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5420618 5',
 mRNA sequence.
 ACCESSION BM042094
 VERSION BM042094.1 GI:16771361
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1CM1875 row: 1 column: 03
 High quality sequence stop: 792.
 Location/Qualifiers
 1..800
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5420618"
 /clone_lib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAC(G). Library constructed by Ling Hong in the

BASE COUNT 146 a 239 c 315 g 100 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,13e-117 Length: 800
 Score: 1203.00 Matches: 244
 Percent Similarity: 99.20% Conservative: 4
 Best Local Similarity: 97.60% Mismatches: 2
 Query Match: 87.55% Indels: 1
 DB: 13 Gaps: 0
 US-09-827-854-15_copy_1_277 (1-277) x BM042094 (1-800)
 QY 1 MetLysValLeuTTPAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 Db 50 ATGAAAGTTCTGTGGGCTCGGTGTGTCACATTCTTGGAGATGCCAGGCAAGGTG 109
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGlnPglInsr 40
 Db 110 GAGCAAGCGGTGAGACAGACCGGAGCCGAGCTGCGCACAGACAGCAGTGGCAGAC 169
 QY 41 GlyGlnArgTTPGluLeuAlaLeuGlyArgPheTTPAspTyrLeuArgTTPValGlnThr 60
 Db 170 GGCACAGCTGGGAACATGGCACGTGGCTTTTGGATTACTGCGCGGTGGAGNA 229
 QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 Db 230 CTGTCTGACAGAGGTGACAGAGAGCTGTCTACCTCCACAGTCAACCAGAACTGAGGCG 289
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 Db 290 CTGATGACAGACACCTGAAGAGTTGAAGCCCTACAAATCCGAATCGAAGAACACTG 349
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 Db 350 ACCCGGTGGCGGAGAGAGACGGGCGCAGCGTGTCCAAAGACACTGAGCGCGCAGGCC 409
 QY 121 ArgLeuGlyAlaAspMetLysAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 Db 410 CGGCTGGGCGGACATGAGAGACGTGTGCGCGCGCTGTGTCAGTACCGCGCGCAGAGTG 469
 QY 141 GlnAlaMetLeuGluGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
 Db 470 CAGGCGCATGCTGGCGGACAGACACCGAGAGCTGGGGTGGCTGGCTCCACACTGGCC 529
 QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 Db 530 AAGCTCGTAAAGCGGCTCTCCGCGATGCGGATGACCTGCAGAAAGCGCTGGCAGTTC 589
 QY 181 GlnAlaGlyAlaArgGluGluGluValArgGlyLysSerAlaIleArgGluArgLeuGly 200
 Db 590 CAGGCGCGGCGCGCGAGGCGCGCGAGCGCGCTCAGCGCCATCCGCGAGGCGCTGGGG 649
 QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPr 220
 Db 650 CCCCCTGTGGAAACAGGCGCGCTGGCGGCGCACACTGTGGCTCCCTGGCGCGGCAAGCC 709
 QY 221 LeuGlnGluArgAlaGlnAlaThrPglGluArgLeuArgAlaArgMetGluGlnMetGly 240
 Db 710 CTACAGAGAGCGGCGCGAGGCTGTGGGCGAGCGGTG-CGGCGCGGATGAGAGAGATGGGC 768
 QY 241 SerArgThrArgAspArgLeuAspGluVal 250
 Db 769 ACGGGAGACCGGCAAGCGCTTGGACAGCGTG 798
 RESULT 5
 LOCUS B1600906 942 bp mRNA linear EST 07-SEP-2001
 DEFINITION 603249241F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301010 5',

Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library.

mRNA sequence.
 BI600906
 VERSION BI600906.1 GI:15493845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 942)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1A11762 row: a column: 11
 High quality sequence stop: 762.
 Location/Qualifiers
 1..942
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5301010"
 /clone_lib="NIH-MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescriptPR (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTCTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to R0T 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."
 BASE COUNT 177 a 284 c 361 g 120 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.62e-117 Length: 942
 Score: 1202.50 Matches: 261
 Percent Similarity: 94.35% Conservative: 6
 Best Local Similarity: 92.23% Mismatches: 10
 Query Match: 87.52% Indels: 7
 DB: 13 Gaps: 1
 US-09-827-854-15_COPY_1_277 (1-277) x BI600906 (1-942)
 Oy 1 MetlysalLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 73 ATGAAGTTCGTGGCGCGCTGCTGTCACATTCGCGCAGATGCGCAGCCAGGTG 132
 Oy 21 GtuglnAlaValGltThgGluProGluProGluLeuArgGlnGlnThgGluTrpGlnSer 40
 |||||
 Db 133 GAGCAAGCGGTGGAGACAGAGCCGCGAGCTGCGCCAGACAGACGAGTGGAGAGC 192
 Oy 41 GlytlnAgtTTPGtLeuAlaLeuGltVArPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 Db 193 GGCACGCCCTGGAGACTGGCTGCTTTTGGATTCCTGCTGGGTGGAGACA 252
 Oy 61 LeuSerGlnGlnValGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 253 CTGTCTGAGCAGGTGCGCAGAGAGCTGCTCAGCTCCAGGTCCACCAAGAACTAGGGCG 312
 Oy 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGluGlnLeu 100
 |||||
 Db 313 CTGATGGAGCAGACCATGAAAGAGTTGAAGGCTTCAAAATCGGAATCGAGGAACTG 372

Oy 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||
 Db 373 ACCCGGTGGCGAGAGACGCGGGCAGGGCTGTCCAGAGAGCTCAGCGCGAGGCC 432
 Oy 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGltTyrArgGlyGluVal 140
 |||||
 Db 433 CGCGTGGCGCGGACATGAGAGACGTGCGGCCCTGCTGTCATACCGCGCGAGGTG 492
 Oy 141 GlnAlaMetLeuGlyGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 493 CAGGCCATGCTCGGCCAAGCAGCAGAGAGCTGCGGGTGGCTGCTCCACTGGCGC 552
 Oy 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 553 AAGCTGCGTAAAGCGGCTCTCCGCGCATGCTGCAGAGAGCGCTGGCACTGATGC 612
 Oy 181 GlnAlaGlyAlaArgGluGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
 |||||
 Db 613 CAGCGCGGGCGCCGCGAGGGCGCCGAGCGCGCTCAGCGCCATCCGCGAGCGCTGGGG 672
 Oy 201 ProLeuValGlnGlnGlyArgValArgValAlaThrPheValGlySerLeuAlaGlyGlnPro 220
 |||||
 Db 673 CCCCTGTGGAAACAGGCGCGCTGCGGCG -GCCACTGTGGCTCCTGCGCGCGAGCGC 731
 Oy 221 LeuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAla -ArgMetGlu-GluMetG 240
 |||||
 Db 732 CTACAGGAGCGGGGCCAGGCGCTGGGGCGAGCGGCTGCGCGCGCGAGTGGAGTGG 791
 Oy 240 LysSerArgThrArgAspArgLeuAsp-GluValLys -GluGlnValAlaGluVal -ArgA 259
 |||||
 Db 792 GCAGCGGAGCAGCGCGCGCTGCGACGAGGAGCAAGCAGGTTGGCGGAAGTTGGCGCG 851
 Oy 259 TalysLeuGlnGluGlnAla---GlnGlnIleArgLeuGlnAlaGlnAlaPheGlnAla 277
 |||||
 Db 852 CCAACTGTGAGAACACAGCGCCAGCAGAAATACGCTTCAGCGAGCGCCTTCAGGCC 910
 RESULT 6
 BI597743
 LOCUS BI597743 922 bp mRNA linear EST 07-SEP-2001
 DEFINITION 60324609F1 NIH-MGC_96 Homo sapiens cDNA clone IMAGE:5300259 5',
 mRNA sequence.
 ACCESSION BI597743
 VERSION BI597743.1 GI:15490682
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 922)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: L1A11760 row: b column: 04
 High quality sequence stop: 782.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5300259"
 /clone_lib="NIH-MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"

QY 240 ySerArgThrArgAsp 245
 |||||
 DB 794 CAGCCGGAACCGCGAA 809

RESULT 9
 BG761746 938 bp mRNA linear EST 15-MAY-2001
 LOCUS 602717942F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:484141 5',
 DEFINITION mRNA sequence.
 ACCESSION BG761746
 VERSION BG761746.1 GI:14072399
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mhc.nci.nih.gov/
 1 (bases 1 to 938)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DC/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCML674 row: c column: 12
 High quality sequence stop: 767.
 Location/Qualifiers
 1. 938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:484141"
 /clone_lib="NIH_MGC_49"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-CDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 library."

BASE COUNT 171 a 272 c 374 g 121 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6,33e-116 Length: 938
 Score: 1187.50 Matches: 265
 Percent Similarity: 93.388 Conservative: 9
 Best Local Similarity: 92.338 Mismatches: 3
 Query Match: 86.438 Indels: 10
 DB: 12 Gaps: 3

US-09-827-854-15_COPY_1_277 (1-277) x BG761746 (1-938)

QY 1 MetLysValLeuThrAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 DB 57 ATCAAGGTTCTGTGGCGGTGTGTCATCTCTCGCAGAGATGCCAGCGCAAGG 116

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
 |||||
 DB 117 GAGCAAGCGGTGGAGACAGACGCCGAGCTCGCCAGACGACGAGTGCGAGAC 176

QY 41 GlyGlnArgTrpGluLeuAlaLeuGluArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 DB 177 GGCACCGCTGGAGACTGGGTGGCTTTGGGATTACTGCGCTGGGTGCAGACA 236

QY 61 LeuSerGluGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 DB 237 CTGTCTGAGCAGGTGCAGGAGAGCTGCTCAGCTCCAGTCCACCGAAGTCCAGCGCG 296

QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGlnGlnLeu 100
 |||||
 DB 297 CTATGTGACGACGACCATGAGAGGATTGAAAGCTTACAAATCGGAATCGAGAACACTG 356

QY 101 ThrProValAlaGlnGluThrThrArgAlaArgLeuSerLysGluLeuGlnAlaGlnAla 120
 |||||
 DB 357 ACCCGGTGGCGAGAGACCGCGGACGCTGTCCAAAGAGCTGCAGCGCGCGAGCGCG 416

QY 121 ArgLeuGluAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyLeuVal 140
 |||||
 DB 417 CGCTGTGGCGCGGACATGAGGACGTGTGGCGCGCTGCTGACATCCGCGCGAGG 476

QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 DB 477 CAGGCAATGCTCGGCCAGACACCGAGAGCTGCGGCTGCGCTCCCTCCACTGCGC 536

QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 DB 537 AAGCTCGTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCTGCGCAGTGTAC 596

QY 181 GlnAlaGlyAlaArgGluGlnGluArgGlyLeuSerAlaLeuArgGluArgLeuGly 200
 |||||
 DB 597 CAGCGCGGGCGCGCAGGAGCGCGCGCGCTGCTCAGCGCCATCCGCGCAGCGCGCTGGGG 656

QY 201 ProLeuValGluGlnGluArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220
 |||||
 DB 657 CCCCTGTGTGACAGGCGCGCGCTGCGCGCGCTGTGGCTCTCCCTTGGCGCGCGCGC 716

QY 220 OleuGlnGluArgAlaGlnAlaAlaTrpGlyGluArgLeuArgAlaArgMetGlu---Glu 238
 |||||
 DB 717 GCTACAGAGAGCGGCG 776

QY 239 MetGlySerArgThrArgAspArgLeuAspGluValLysGluGlnValAla---Glu 256
 |||||
 DB 777 TGGGCG 836

QY 257 ValArgAlaLysLeuGlnGluGlnGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGln 275
 |||||
 DB 837 TTGGCGCGCAAGCTGGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896

QY 275 eGlnAla 277
 |||||
 DB 897 CCAGCGCC 903

RESULT 10
 BG829472 845 bp mRNA linear EST 22-MAY-2001
 LOCUS 602763768F1 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:4899112 5',
 DEFINITION mRNA sequence.
 ACCESSION BG829472
 VERSION BG829472.1 GI:14177059
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mhc.nci.nih.gov/
 1 (bases 1 to 845)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1CM1790 row: g column: 17
High quality sequence stop: 829.

FEATURES

Location/Qualifiers

1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:489112"
/clone_lib="NIH_MGC_42"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 153 a 244 c 342 g 106 t

ORIGIN

Alignment Scores:
Pred. No.: 1.03e-114 Length: 845
Score: 1175.50 Matches: 253
Percent Similarity: 96.23% Conservative: 2
Best Local Similarity: 95.47% Mismatches: 7
Query Match: 85.55% Indels: 5
DB: 12 Gaps: 2

US-09-827-854-15_COPY_1_277 (1-277) x BG829472 (1-845)

QY 1 MettysValleuTPRALalaleuValThrpheleuAglCysGlnAlaIysVal 20
DB 57 ATGAAGGTTCTGTGGCTGCGTTCGTGTCATTCCTGGCAGATGCCAGGCAAGG 116
QY 21 GlnGlnAlaValGluThrGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
DB 117 GAGCAAGCGGTGGAGACAGAGCGCGGCCGAGCTGCCGACAGACGAGGAGG 176
QY 41 GlnGlnAgtTrpGluLeuAlaLeuGlyArgPheTrpAspTrpLeuArgTrpValGlnThr 60
DB 177 GGCACGCGCTGGGAACCTGCTGCTGCTTTGGGATTAACCTCGCTGGGTGCAGACA 236
QY 61 LeuSerGlnGlnValGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
DB 237 CTGTGTGACAGAGTGCAGAGAGAGCTGCTCACTCCAGGTACCCAGAACTGAGGGCG 296
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGluLeu 100
DB 297 CTGATGAGACGAGACCATGAGAGAGTTGAAGCGCTCAAAATGGAAGTGGAGACACACTG 356
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaIaGlnAla 120
DB 357 ACCCGGTTGGCGAGAGAGCGCGGCGCTGTCCAGAGAGCTCCAGCGCGCGCAGAGCC 416
QY 121 ArgLeuGlyAlaAspMetGlnAspValCysGlyArgLeuValGlnTrpArgGlyGluVal 140
DB 417 CGGCTGGCGCGGACATGAGAGAGCTGTGCGCGCTGTGTCAGTACCGCGCGAGGGG 476
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
DB 477 CAGGCGCATGCTCGCGCAGAGCACCGAGAGCTGGGGGTGGCGCTGCCACTGGCGC 536
QY 161 LysLeuAgtLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValLys 180
DB 537 AAGCTGCGTAACGGGCTCTCCGCGATGCGCATGACCTGCAGAAAGCGCTGCGCACTGTAC 596
QY 181 GlnAlaGlyAlaArgGlnGlyValArgLysArgLysSerAlaIlaArgGluArgLeuGly 200
DB 597 CAGGCGCGGGCGCGCGAGGGCGCGCAGCGCGCTCCAGCGCCATCCGAGAGCGCTGGGG 656
QY 201 ProLeuValGlnGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyGlnPro 220

DB 657 CCCCTGTGGAGAACAGAGCGCGCTGCGGGCGCG -ACTGTGGGCTCCCTTGGCGGAG -CCG 714
QY 221 LeuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGlnGluMetGly 240
DB 715 CTACAGGAGCGGGCGCAGCGTGTG---GGGCGAGCGGTGCCCGCGGATGAGAGATGGGC 771
QY 241 SerAgtTrp-ArgAspArgLeuAspGluValLysGlnGlnValAlaGlnValArgAlaGly 260
DB 772 AGCCGAGCCGCCA---CGCTGTGACGAGGTGAAGACAGGTGGCGAGGTGCCGCCA 828
QY 260 sLeuGlnGluGln 264
DB 829 GCTTGAGAGAGCAG 841

RESULT 11

LOCUS BG472299 927 bp mRNA linear EST 21-MAR-2001

DEFINITION 602513830F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4645759 5', mRNA sequence.

ACCESSION BG472299
VERSION BG472299.1 GI:13404485

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 927)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.llnl.gov
Plate: L1CM1419 row: k column: 08
High quality sequence stop: 848.

FEATURES

source

Location/Qualifiers

1..927
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645759"
/clone_lib="NIH_MGC_16"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site:1: XhoI; Site:2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 170 a 271 c 371 g 115 t

ORIGIN

Alignment Scores:
Pred. No.: 3.53e-114 Length: 927
Score: 1171.00 Matches: 262
Percent Similarity: 93.93% Conservative: 1
Best Local Similarity: 93.57% Mismatches: 10
Query Match: 85.23% Indels: 9
DB: 12 Gaps: 0

US-09-827-854-15_COPY_1_277 (1-277) x BG472299 (1-927)

QY 1 MettysValleuTPRALalaleuValThrpheleuAglCysGlnAlaIysVal 20
DB 47 ATGAAGGTTCTGTGGCTGCGTTCGTGTCATTCCTGGCAGATGCCAGGCAAGGTG 106

QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
 |||||
 Db 107 GAGCAACGGGTGGAGACAGAGCCGAGCCGAGCTGCGCAGCAGACCGAGTGGCAGAGC 166
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 Db 167 GGCACGGCTGGAGACTGGACTGGCTTTGGGATTTACTGCTGGCTGGGTGCAGACA 226
 QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 227 CTGTCTGAGCAGAGTGCAGAGAGAGCTCTCAGCTCCAGAGTCAACCGAAGTGAAGGGCG 286
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||
 Db 287 CTGTATGACAGACACATGAGAGAGTTGAGAGCTTCAAAATCGGAACGTGAGAACTG 346
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 347 ACCCGGTGGCGGAGAGAGCGGGCAGCTGTCCAAAGAGCTGCAGCGCGCCAGGCC 406
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||
 Db 407 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGGTGCAGTACCGCGGAGAGTG 466
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 467 CAGCGCATGCTCGCGCCAGACACCGAGAGCTGCGGCTGCGCTCCGCTCCACTCGCGC 526
 QY 161 LysLeuArgLysArgLeuArgLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 527 AAGCTGTGTAAGGAGGCTCTCCCGGATGCCGATGACTGCAAGAGCCCTGGCAGTGTAC 586
 QY 181 Gln-AlaGlyAlaArgGluGluAlaGluArgGly--LeuSerAlaIleArgGluArgLeu 199
 |||||
 Db 587 CAGGCGCGGGGCGCGGAGGGCGCGGCGCCCTCAAGCGCCATCCGCGAGCGGCTT 646
 QY 200 --GlyProLeuValGluGlnGlyArgValArgAlaAlaIleThrValGlySerLeuAlaGly 219
 |||||
 Db 647 GGGGCCCTCTGGTGAACAGAGCGCGCTGCGGCGCCGCTGAGCTGCTCCCTGCGCGCC 706
 QY 219 LnpProLeuGlnGluArgAlaGlnAlaIleThrValGluArgLeuArgAlaArgMetGluGln 239
 |||||
 Db 707 AGCG-CTACAGAGAGCGGGCCAGCGCTGGGGCCAGACGGTG-CGCGCGCGGATGAGAGAGA 764
 QY 239 etGlySerArg-ThrArgAspArgLeuAspGluValLysGluGlnValAlaGlu-ValAr 258
 |||||
 Db 765 TGGGGCAGCGGGAGACCGCGCAACGCTGACCGAGGTGAAGAGACAGTGGCGGAGGTGCG 824
 QY 258 GAlaLysLeuGluGluGlnAlaGlnGlnIleArgLeuGlnAlaGlu 273
 |||||
 Db 825 CCGCAAGCTGAGAGAAAGCAGCGCAGAGATGAGGCTGACAGCGCGAG 870
 RESULT 12
 BG707147 790 bp mRNA linear EST 07-MAY-2001
 LOCUS BG707147
 DEFINITION 602670283F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793084 5',
 mRNA sequence.
 ACCESSION BG707147
 VERSION BG707147.1 GI:13983201
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 790)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LLAM10672 row: e column: 21
 High quality sequence stop: 781.
 Location/Qualifiers
 1. 790
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4793084"
 /clone_lib="NIH_MGC_96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3',
 size selected for average insert size 2.3 kb and
 normalized to ROT 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."
 BASE COUNT 143 a 234 c 311 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,94e-112 Length: 790
 Score: 1152.00 Matches: 236
 Percent Similarity: 98.74% Conservative: 0
 Best Local Similarity: 98.74% Mismatches: 3
 Query Match: 83.84% Indels: 1
 DB: 12 Gaps: 0
 US-09-827-854-15_COPY_1_277 (1-277) x BG707147 (1-790)
 QY 1 MetLysValLeuThrPheAlaLeuLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
 |||||
 Db 73 ATGAGGTTCTGTGGGCTGTGCTGTGTCATTCCTGTCAGAGATGCCAGGCCAAGGTG 132
 QY 21 GluGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluThrGlnSer 40
 |||||
 Db 133 GAGCAACGGGTGGAGACAGAGCCGAGCGCAGCTCGCCAGAGACCGAGTGGCAGAGC 192
 QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
 |||||
 Db 193 GGCACGGCTGGAGACTGGACTGGCTTTGGGATTTACTGCTGGGTGCAGACA 252
 QY 61 LeuSerGluGlnValGlnGluLeuLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
 |||||
 Db 253 CTGTCTGAGCAGAGTGCAGAGAGAGCTGCTCAGCTCCAGAGTCAACCGAAGTGAAGGGCG 312
 QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
 |||||
 Db 313 CTGTATGACAGACACATGAGAGAGTTGAGAGCTTCAAAATCGGAACGTGAGAACTG 372
 QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
 |||||
 Db 373 ACCCGGTGGCGGAGAGAGAGCGGGCAGCTGTCCAGAGTGCAGCGCGCGAGGCC 432
 QY 121 ArgLeuGlnAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
 |||||
 Db 433 CGGCTGGCGCGGACATGAGAGAGCTGTGGCGCCCTGTGGTGCAGTACCGCGGAGAGTG 492
 QY 141 GlnAlaMetLeuGlyGlnSerThrGluLeuArgValArgLeuAlaSerHisLeuArg 160
 |||||
 Db 493 CAGGCAATGCTCGCGCAGAGAGAGCTGCGGGTCCGCTCCCTCCCACTGCGCC 552
 QY 161 LysLeuArgLysArgLeuArgLeuArgAspAlaAspLeuGlnLysArgLeuAlaValTyr 180
 |||||
 Db 553 AAGCTGTGTAAGGAGGCTCTCCCGATGCGGATGACTGCAAGAGCGCTGGCAGTGTAC 612

QY 181 GlnAlaGlyAlaArgLuglAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGCGCGGGCCCGCAGGCGCGCGGCTCAGCGCCATCCGAGCGGCTCGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
|||||
Db 673 CCCCCTGGTGAACAGGCGCGCGCGCCACTGTGGCTCCCTCGGCAGCGCGCG 732
QY 221 LeuGlnGluArgAlaGlnAlaThrpglYgluArgLeuArgAlaArgMetGluGlnMet 239
|||||
Db 733 CTACAGGACCGGCGCCAGGCTGGGCGGAGCGGTG-CGCGCGGATGAGAGATG 788
RESULT 13
B1551475 919 bp mRNA linear EST 05-SEP-2001
LOCUS 603194314F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:5274003 5',
DEFINITION mRNA sequence.
ACCESSION B1551475
VERSION B1551475.1 GI:15438787
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 919)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
http://image.llnl.gov
Plate: LLM11691 row: 1 column: 04
High quality sequence stop: 812.
Location/Qualifiers
1..919
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5274003"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size selected for average insert size 2.5 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NHGRI/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 172 a 270 c 363 g 114 t
ORIGIN
Alignment Scores:
Pred. No.: 7.61e-112 Length: 919
Score: 1149.00 Matches: 247
Percent Similarity: 93.63% Conservative: 3
Best Local Similarity: 92.51% Mismatches: 16
Query Match: 83.62% Indels: 4
DB: 13 Gaps: 0
US-09-827-854-15_COPY_1_277 (1-277) x B1551475 (1-919)
QY 1 MetLysValLeuTrpAlaAlaLeuValThrPheLeuAlaGlyCysGlnAlaLysVal 20
|||||

Db 73 ATGAAGTTCTGTGGCGTCGCTTGCTGTGTCACATTCTGGCAGATGCCAGGCAAGTG 132
QY 21 GluGlnAlaValaGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
|||||
Db 133 GAGCAAGCGGTGGAGACAGAGCCGAGCCGAGCTGCGCCACACAGACCGATGGCAGAGC 192
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlnArgPheTrpAspPyrLeuArgTrpValGlnThr 60
|||||
Db 193 GGCACAGCGCTGGAACTGGACATGGGTGGCTTTTGGATATTACTCGCGGTGGTGCACA 252
QY 61 LeuSerGluGlnValaGlnGlnGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
|||||
Db 253 CTGTGTGACACAGTGCAGAGAGAGCTGCACCTCCAGGCTCACCCAGGAACCTGAGGGCG 312
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTrpLysSerGluLeuGlnGlnGlnLeu 100
|||||
Db 313 CTGATGAGACAGACCATGAGAGAGTTGAAGGCTCAATAATCGAACTGGAGAACACTG 372
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGluLeuGlnAlaAlaGlnAla 120
|||||
Db 373 ACCCGGTGGCGGAGAGAGAGCGGCGACGCTGTCCAGGAGACTGCAGCGCGCGAGCGC 432
QY 121 ArgLeuGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTrpArgGlyGlnVal 140
|||||
Db 433 CGGCTGGGCGCGACATGAGAGAGCTGTGCGCGCCCTGTGTGCTACCGCGCGAGGTG 492
QY 141 GlnAlaMetLeuGlnGlnSerThrGlnGluLeuArgValArgLeuAlaSerHisLeuArg 160
|||||
Db 493 CAGGCGCATGCTGGCCAGAGACCGAGAGCTGGCGGCTGCGCTCCCTCCCACTGGCGC 552
QY 161 LysLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
|||||
Db 553 AAGCTGCGTAACCGGCTCTCCGCGATGCCATGACCTGACCAACAGCGCTGCGCATGTAC 612
QY 181 GlnAlaGlyAlaArgGlnGlyAlaGluArgGlyLeuSerAlaIleArgGluArgLeuGly 200
|||||
Db 613 CAGGCGCGGGCGCGATGGCGCCAGCGCGTGTGACGCGCATCGCGAGCGCGCTGGGG 672
QY 201 ProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyInPro 220
|||||
Db 673 CCCCCTGGTGAACAGGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGAGCCCT 732
QY 221 LeuGlnGluArgAlaGlnAlaThrpglYgluArgLeuArgAlaArgMetGluGlnMet 240
|||||
Db 733 A--CAGAGAGCGGCGCAGGCGCTGGGGCGAGCGGTGCGCGGATGAGAGAAATGGG 790
QY 240 ySerArgThrArgAspArgLeuAspGluValLysGlnGlnValaGluValaArgAlaLys 260
|||||
Db 791 CAGCGCGAGCCCGGAGCG--CTGGACGAAGTGAAGAGCACGTGGCGGAGGG-TGCGCCAA 847
QY 260 sLeuGlnGluGlnAlaGln 266
|||||
Db 848 GCTGAGGACCGAGCCAGAG 866
RESULT 14
BM042228 757 bp mRNA linear EST 07-NOV-2001
LOCUS 603616186F1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:5557004 5',
DEFINITION mRNA sequence.
ACCESSION BM042228
VERSION BM042228.1 GI:16771495
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS 1 (bases 1 to 757)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: DCTD/DTP


```
QY 21 GIuGlnAlaValGluThrGluProGluProGluLeuArgGlnGlnThrGluTrpGlnSer 40
    |||||
Db 135 GAGCAAGCGGTGGACAGACAGCGGAGCCGAGCTGCCGCCACAGACAGCCAGTGGCAGAGC 194
    |||||
QY 41 GlyGlnArgTrpGluLeuAlaLeuGlyArgPheTrpAspTyrLeuArgTrpValGlnThr 60
    |||||
Db 195 GGCACGCGTGGAACTGGCTGCTTTGGGATTACCTGCGCTGGGTGCAGACA 254
    |||||
QY 61 LeuSerGluGlnValGluGluLeuLeuSerSerGlnValThrGlnGluLeuArgAla 80
    |||||
Db 255 CTGTCTGACAGAGTCAGAGAGAGCTGCTCAGCTCCAGGTCACCCAGAACTGAGGGCG 314
    |||||
QY 81 LeuMetAspGluThrMetLysGluLeuLysAlaTyrLysSerGluLeuGluGlnLeu 100
    |||||
Db 315 CTGATGAGACAGACCATGATGAGAGTTGAAGCGCTTACAAATCGAACTGAGAGACACTG 374
    |||||
QY 101 ThrProValAlaGluGluThrArgAlaArgLeuSerLysGlu-LeuGlnAlaAlaGlnAl 120
    |||||
Db 375 ACCCGGTGGCGGAGAGACGCGGCGCTGCTCAAGGAAGCTGACAGCGGCGCAGAGC 434
    |||||
QY 120 ArgGluGlyAlaAspMetGluAspValCysGlyArgLeuValGlnTyrArgGlyGluVal 140
    |||||
Db 435 CCGGCTGGGCGGACATGGAGAGCTGTGGCGCCCTGGGTGACGTACCGCGCGAGAGT 494
    |||||
QY 140 GlnAlaMetLeuGlnSerThrGluGluLeuArgValArgLeuAlaSerHisLeuArg 160
    |||||
Db 495 GCAGGCCATGCTCGGCCAGAGACCGAGAGAGTGGCGCTGCCCTCCACCTCGC 554
    |||||
QY 160 GlyLeuArgLysArgLeuLeuArgAspAlaAspAspLeuGlnLysArgLeuAlaValTyr 180
    |||||
Db 555 CAAGCTGCGTAAGCGGCTCTCCGCGATGCCGATGACCTGCAGAAAGCGCCTGGCAGTGTA 614
    |||||
QY 180 GlnAlaGlyAlaArgGluGluGluArgGlyLeuSerAlaIleArgGluArgLeu-G 200
    |||||
Db 615 CCAGGCCGCGGCCCGGAGGCGCGGAGCGGCGCTCAGCGCCATCCCGAGCGCCTGGG 674
    |||||
QY 200 LysProLeuValGluGlnGlyArgValArgAlaAlaThrValGlySerLeuAlaGlyIle 220
    |||||
Db 675 GCCCTTGGTGGACAGAGGCGCGCTGCCGCGGCCACCTGTGGGCTCCCTGNGCGGCAGC 734
    |||||
QY 220 IsoleuGlnGluArgAlaGlnAlaTrpGlyLysArgLeuArgAlaArgMetGluGluMetG 240
    |||||
Db 735 CGCTACAGAGAGCGGCGCAGGCTG---GGCGCAGCGGTGCGCGCGGATGGAGGAGATGG 791
    |||||
QY 240 LysArgTrpThrArgAspArgLeuAspGluValLysGluGlnValAlaGluValArg 258
    |||||
Db 792 GCAG-CGGACCGG---ACGCTGAGCAGAGTG-AAAGAGCAGTGGGGAAGTGTGCCG 842
    |||||
```

Search completed: March 14, 2003, 20:13:55
Job time : 1129.96 secs

